DOCUMENT RESUME

ED 064 149 SE 014 128

AUTHOR	Higgins, Jon L.
TITLE	Research Reporting Sections, National Council of Teachers of Mathematics 50th Annual Meeting, Chicago, Illinois, April 16-19, 1972.
INSTITUTION	ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, Columbus, Ohio.
PUB DATE	Apr 72
NOTE	78p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Achievement; *Conference Reports; Elementary School Mathematics; *Instruction; *Learning; *Mathematics Education; *Research; Secondary School Mathematics; Teacher Education

ABSTRACT

This is a collection of 20 abstracts of research papers presented at the 50th annual meeting of the National Council of Teachers of Mathematics. Four papers were presented at each of five research reporting sections: (1) mathematics instruction and instructional materials, (2) aspects of mathematics learning, (3) general research topics in mathematics education, (4) mathematics achievement and its correlates, and (5) teacher education and evaluation. Topics covered include the teaching of division, CAI, feedback, solving open sentences, learning strategies, studying mathematics texts, negations and problem solving, low achievers, addition and subtraction, remedial mathematics in the junior college, inquiry-centered teaching, mastery learning at the elementary level, mathematics teachers' learning strategies, teacher education, videotapes and audiotapes in questioning, and evaluation of elementary school teachers. (JM)

ED 064147

[redacted]

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SMEAC/SCIENCE, MATHEMATICS, AND ENVIRONMENTAL EDUCATION INFORMATION ANALYSIS CENTER

. . . an information center to organize and disseminate information and materials on science, mathematics, and environmental education to teachers, administrators, supervisors, researchers, and the public. A joint project of The Ohio State University and the Educational Resources Information Center of USOE.

MATHEMATICS EDUCATION REPORTS

RESEARCH REPORTING SECTIONS
NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS
50th ANNUAL MEETING

Chicago, Illinois
April 16-19, 1972

ERIC Information Analysis Center for
Science, Mathematics, and Environmental Education
1460 West Lane Avenue
Columbus, Ohio

PREFACE

The ERIC Information Analysis Center for Science, Mathematics, and Environmental Education has compiled abstracts of the research papers to be presented at the 50th annual meeting of the National Council of Teachers of Mathematics. Minor editing has been done by the ERIC staff to provide a general format for the papers. Many of the papers that are abstracted here will be published in journals or be made available through the ERIC system. These will be announced in *Research in Education* and in the journal *Investigations in Mathematics Education*.

April, 1972

Jon L. Higgins
Associate Director for
Mathematics Education

Sponsored by the Educational Resources Information Center of the United States Office of Education and the Center for Science and Mathematics Education, The Ohio State University.

This publication was prepared pursuant to a contract with the Office of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgement in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

CONTENTS

NCTM RESEARCH REPORTING SECTIONS
1972 Annual Meeting

Richard Kratzer
A COMPARISON OF INITIALLY TEACHING DIVISION EMPLOYING THE DISTRIBUTIVE AND GREENWOOD ALGORITHM WITH THE AID OF A MANIPULATIVE MATERIAL 2

Wayne H. Hall
THE EFFECT ON PERFORMANCE OF FEEDBACK, NUMBER OF EXERCISES, AND AMOUNT OF DETAIL 5

Harold L. Schoen
A COMPARISON OF FOUR TYPES OF FEEDBACK TO STUDENT RESPONSES IN A CAI PROGRAM DESIGNED TO TEACH THE CONCEPT OF FUNCTION 10

Thomas E. Foster
THE EFFECT OF COMPUTER PROGRAMMING EXPERIENCES ON STUDENT PROBLEM SOLVING BEHAVIORS IN EIGHTH GRADE MATHEMATICS 12

Douglas A. Grouws
METHODS OF SOLUTION USED IN SOLVING OPEN SENTENCES 16

Lloyd F. Scott
CHILDREN'S NUMBER PREFERENCES: GEOMETRY VS. ABSTRACT . . . 19

Schmuel Avital
LEARNING OF CONJUNCTIVE CONCEPTS IN MATHEMATICS IN THE JUNIOR HIGH SCHOOL 22

A. B. Durell
STRATEGIES FOR LEARNING MATHEMATICAL CONCEPTS 25

Alan W. Holz
A TECHNIQUE FOR STUDYING THE ORGANIZATION OF MATHEMATICS TEXTS 28

Gerald Kulm
NEGATIONS IN LINEAR SYLLOGISMS AND PROBLEM SOLVING 32

Constance Anne Carroll
LOW ACHIEVERS' UNDERSTANDING OF FOUR LOGICAL INFERENCE FORMS: AN ANALYSIS OF DIFFICULTIES AND OF THE EFFECT OF INSTRUCTION 35

Susan Rudder Fuller
A STUDY OF PROBLEM SOLVING METHODS OF STUDENTS WITH AND WITHOUT THE CONSTRAINT OF A TIME LIMIT 39

Larry E. Wheeler
THE RELATIONSHIP OF MULTIPLE EMBODIMENTS OF THE REGROUPING CONCEPT TO CHILDREN'S PERFORMANCE IN SOLVING MULTI-DIGIT ADDITION AND SUBTRACTION EXAMPLES . . . 42

Ronald R. Edwards
THE PREDICTION OF SUCCESS IN REMEDIAL MATHEMATICS COURSES IN THE PUBLIC COMMUNITY JUNIOR COLLEGE 45

Irving M. Brauer
THE EFFECTS OF VARYING DEGREES OF INQUIRY-CENTERED TEACHING ON COGNITIVE ACHIEVEMENT IN MATHEMATICS UNDER DIFFERENT SCHOOL CONDITIONS 52

Mildred E. Kersh
A STUDY OF MASTERY LEARNING IN ELEMENTARY MATHEMATICS . . . 56

Nicholas A. Branca
AN EXPLORATORY STUDY OF EXPERIENCED MATHEMATICS TEACHERS' STRATEGIES IN LEARNING A MATHEMATICAL STRUCTURE 62

Clinton A. Erb
A FORMATIVE EVALUATION OF AN EXPERIMENTAL TEACHER EDUCATION PROJECT FOR JUNIORS IN MATHEMATICS EDUCATION . . 64

Ellen L. Bortz
THE EFFECTIVENESS OF VIDEOTAPE AND AUDIOTAPE IN IMPROVING THE COGNITIVE LEVEL OF QUESTIONS IN THE MATHEMATICS CLASSROOM 67

W. D. McKillip and M. Mahaffey
EVALUATING ELEMENTARY SCHOOL TEACHERS KNOWLEDGE OF THE CONTENT AND METHODS OF TEACHING ELEMENTARY SCHOOL MATHEMATICS 69

Monday, April 17, 1972

12:30 p.m. - 2:00 p.m. Bel Air Room

Research Reporting Section
NCTM Section No. 29

Subject: Mathematics Instruction and Instructional Materials

Presider: Robert E. Reys, University of Missouri, Columbia, Missouri

Speakers: Richard Kratzer, University of Maine--Gorham
A COMPARISON OF INITIALLY TEACHING DIVISION EMPLOYING THE DISTRIBUTIVE AND GREENWOOD ALGORITHM WITH THE AID OF A MANIPULATIVE MATERIAL

Wayne H. Hall, Eastern Washington State College
THE EFFECT ON PERFORMANCE OF FEEDBACK, NUMBER OF EXERCISES, AND AMOUNT OF DETAIL

Harold L. Schoen, Virginia Polytechnic Institute and State University
A COMPARISON OF FOUR TYPES OF FEEDBACK TO STUDENT RESPONSES IN A CAI PROGRAM DESIGNED TO TEACH THE CONCEPT OF FUNCTION

Thomas E. Foster, Sam Houston State University
THE EFFECT OF COMPUTER PROGRAMMING EXPERIENCES ON STUDENT PROBLEM SOLVING BEHAVIORS IN EIGHTH GRADE MATHEMATICS

Reactor: Jeremy Kilpatrick, Teachers College, Columbia University

2

A COMPARISON OF INITIALLY TEACHING DIVISION EMPLOYING THE DISTRIBUTIVE AND GREENWOOD ALGORITHM WITH THE AID OF A MANIPULATIVE MATERIAL

Richard Kratzner
University of Maine--Gorham
Gorham, Maine 04038

Most elementary mathematics texts develop long division by the method of repetitive subtraction involving the Greenwood algorithm. This current trend in developing long division centers around Van Engen and Gibb's study General Mental Functions Associated with Division.[1] Also, most current texts employ a transition stage from the use of the Greenwood algorithm to the traditional distributive algorithm.

The purpose of this study was again to raise the question whether the use of the Greenwood algorithm to teach division of whole numbers is more effective than the use of the distributive algorithm and to keep in mind most current texts make the above transition.

In this study, two instructional units were prepared where both methods were taught in a meaningful manner by recording manipulation of popsicle sticks. One method used a partitioning rationale and the distributive algorithm as a method of keeping records of manipulating bundles of sticks; the other used a subtractive rationale and the Greenwood algorithm as a method of keeping records of manipulating bundles of sticks.

Two schools were selected in which each contained six fourth grade self-contained classrooms and used the same elementary arithmetic text. The teachers in each school were responsible for all subjects except art, music and physical education. Previous to the study, both schools had employed the manipulative materials, which were involved in the presentation of each division approach, in teaching place value. Before the instructional programs were started, the two division approaches were randomly assigned to each of three classes in each school. Each teacher involved in the study was given instruction in the assigned approach; a guide was also provided containing the instructions and the worksheets used in developing that particular approach. Multiplication tables consisting of facts through ten were supplied to each student.

Two tests were administered immediately at the completion of the sequence, administered four weeks later as a measure of retention and then administered the following fall as a measure of delayed retention. The first test, employed by Van Engen and Gibb, consisted of sixteen division examples (one, two and three-digit divisors). The second test employed fourteen verbal problems. The test examples were classified by the participating teachers as familiar--taught in the sequence, and unfamiliar--not taught in the sequence. The ability of a student to achieve on unfamiliar problems was his measure of understanding of the division process. As Van Engen and Gibb summarized in their study about the subtractive methods group, "...the subtractive methods group had not reached a high level of skill in solving familiar problems yet their understanding of the process was such that they were better able to transfer to new situations."[1]

---

[1] H. Van Engen and G. Gibb, General Mental Functions Associated with Division (Issue II: Cedar Falls, Iowa: Iowa State Teachers College Educational Service, 1956).

The class means (N = 12) of the two instructional groups on the immediate test, the retention test and the delayed retention test was examined by using analysis of covariance. The five per cent level of significance was adopted.

The first hypothesis--there is no difference between the two experimental methods on a test of computational problems--was analyzed three ways, that is by using the results of the total set of computational problems, by using the results of the computational problems that were classified as familiar and by using the results of the computational problems that were classified as unfamiliar. There existed some problems in neither classification. The computational problems were scored each of two ways, by giving partial credit and by right-wrong. In the final analysis, the significance of the hypothesis was not affected by the method of scoring and the direction of the difference of the means was always the same. The covariates considered for the first hypothesis were class mean scores of intelligence as measured by the Kuhlmann Finch Test (subtractive - 109.4, distributive - 111.2) and class mean scores of arithmetic computation as measured by the Stanford Achievement Test (subtractive - 3.8, distributive - 3.7).

The second hypothesis--there is no difference between the two experimental methods on a test of verbal problems--was analyzed the above three ways. The verbal problems were classified by the rationale. There existed one problem in neither classification. The covariates considered for the second hypothesis included the above two and the class mean scores of paragraph meaning (subtractive - 4.4, distributive - 4.2).

The first hypothesis was rejected for the total set of computational problems in favor of the distributive instructional group on the immediate test, retention test and the delayed retention test. The method of scoring did not affect the significance of the hypothesis. The distributive group average based on right-wrong scoring was greater than the subtractive approach with partial credit scoring if two points rather than one point were allowed for each correct answer on the distributive group scoring. The distributive group average on the delayed retention test was greater than the subtractive group average on the immediate test.

---

[1] Van Engen and Gibb

The first hypothesis was considered for familiar computational problems. The null hypothesis was not rejected for either scoring procedure on the immediate test, on the retention test or the delayed retention test. The two point convention and the average increase for the distributive group above holds in this situation.

The hypothesis was then considered for those problems that were classified as unfamiliar computational problems. The null hypothesis was rejected in each case (immediate, retention and delayed retention testing) at the 0.05 level of significance with the direction of the difference favoring the distributive group. It should be noted that the subtractive group increased in achievement on the retention test as compared to the immediate test. The results of the analysis for the distributive group yielded the same conclusions as that for the total set of problems.

The second hypothesis was first analyzed using the scores obtained from the total set of verbal problems. The hypothesis was not rejected on the immediate test results but was rejected for the retention test and the delayed retention test. The direction of the difference favored the distributive group. Again the findings were the distributive group average increased over the first retention period and their average on the delayed retention test was greater than the average of the subtractive group on the immediate test.

The following summarizes the results as found in the domain of this study:

1. There was no difference between the distributive and subtractive approaches of teaching long division in achievement on problems similar to those problems studied within the sequence. If problems studied within the sequence is the main concern, either instructional method may be considered in teaching division of whole numbers.

2. There was a significant difference between the distributive and the subtractive approaches of teaching long division in achievement on computational problems not involved within the instructional sequence, that is on unfamiliar problems. The distributive group better understood the division process.

It would seem, in conclusion, within the domain of this study if the distributive algorithm is to be the terminal algorithm of the division sequence, then the distributive algorithm with meaning should be considered in developing division of whole numbers.

THE EFFECT ON PERFORMANCE OF FEEDBACK, NUMBER OF EXERCISES, AND AMOUNT OF DETAIL

Wayne H. Hall
Eastern Washington State College
Cheney, Washington 99004

The idea that learning is improved when a reinforcer or some feedback promptly follows one's behavior is found in such prominent theories of learning as those of Hull (1952), Skinner (1953), and Spence (1956). Many textbooks in psychology and educational psychology mention the superiority of immediate feedback as a principle of human learning and point out its presumed importance in educational practices. Also, many educators have come to feel that reinforcement should be considered as a major, and perhaps a necessary, condition for most learning.

Even though it is felt by many educators that reinforcement should be considered as a major condition for most learning, this view is evidenced in very few mathematics textbooks. One finds that some of the textbooks have some answers in the back of the book, others have no answers and some have a separate answer book. However, the use of answers, if available, is left to the discretion of the individual teacher. Thus, it would seem that some research in the field of reinforcement theory relative to the study of mathematics would be of benefit to the teacher when he attampts to decide on the use or disuse of answers.

A few experiments have been devised which studied the effects of different schedules of reinforcement. Such studies by Walbesser (1964), Krumboltz and Weisman (1962), and Oppenheimer (1968) failed to find any differential effects attributable to the different schedules of reinforcement considered. Not only did they fail to find any difference in schedules, but they also failed to find any differences favoring reinforcement over nonreinforcement. However, studies by Chansky (1960), Bourne and Pendleton (1958), and Bourne and Haygood (1960), have found continuous reinforcement to be superior to various schedules of intermittent reinforcement in the acquisition of concepts.

An idea which, in the field of mathematics, might be considered as closely allied to the concept of reinforcement is that practice is a necessary part of the learning process. DeCecco (1968) points out that unless other learning conditions are ideally provided, we probably learn very little (of relative permanency) from the first response we make to a particular stimulus.

The importance of practice has been studied by several researchers. Maertens (1969) found no significant differences in the achievement of students doing homework as compared to those not doing homework. Also, studies by Gagné (1965) and Bassler, Curry, Hall, and Mealy (1969) found no significant differences in student achievement relative to a number of exercises variable. In contrast to the preceding findings, Koch (1965) found that experimental groups doing daily homework assignments achieved significantly higher scores on tests of computational skill than did those doing no homework.

Another concept which seems of interest relative to the above two concepts is that of detail. That is, if answers are to be supplied, which is more beneficial: answers worked out in their entirety or answers with no work shown?

Very few studies have been conducted in which the effects of detail were closely scrutinized. However, Gitman (1967) conducted a study in which he compared the effects on achievement of response contingent feedback and simply knowledge of the correct result. He found no differences in achievement relative to the feedback dimension.

Thus, it would seem that if relationships exist among the three variables mentioned above, such relationships might have a significant bearing on classroom procedures. Such procedures might be altered on the basis of a teacher's knowledge of existing relationships. Also, if important relationships exist among the three variables, publishers of textbooks should be aware of such relationships so they could, perhaps, incorporate such knowledge into their publication of textbooks.

The specific purposes of this study were to investigate (1) if a difference in performance existed among students working 5, 15, or 30 practice exercises each day for 4 days; (2) if differences in performance existed among students receiving no feedback, partial (1/2) feedback, or continuous feedback; (3) if differences in performance existed for students receiving answers in the form of problems worked in detail as opposed to those receiving only answers; (4) the relevant interactions (or relationships) existing among the above described variables.

The study was conducted in a situation which closely approximated actual classroom conditions. That is, the experimental treatments were administered in the setting of the mathematics classroom with the classroom teacher presiding over the class. Also, the variables, though somewhat global in nature, were construed as being variables which are encountered in actual classroom situations and which could be controlled, to some extent, in their normal setting.

The experimental treatments consisted of a number of addition of fractions problems and each treatment reflected different numbers of exercises (5, 15, 30), different contingencies of feedback (none, partial (1/2), continuous), and different amounts of detail (none, full). Feedback consisted of the answers being made available to the subjects, and detail meant that the answers were worked in their entirety.

The experimental treatments were presented in the form of 8 1/2 inch by 5 1/2 inch booklets prepared by the experimenter. There was one problem per page and answers, where utilized, were presented on the back of the page in such a manner that the student could fold the page over to compare the correct answer with his answer. The total number of problems (addition of fractions and filler items) to be worked each day was 36. As the number of addition of fractions problems decreased, the number of filler items increased.

The experiment consisted of four days of treatment which was administered on Tuesday and Thursday of two consecutive weeks. The content area in which the study was conducted was a review of addition of fractions at the sixth-grade level.

The sample space on which the final analysis was conducted consisted of 165 subjects. These subjects were members of seven intact classes and were randomly distributed throughout 15 experimental groups with nine subjects in each group and two control groups of 15 subjects each.

The criterion measure consisted of a 41-item pre- and post-test, constructed by the experimenter. The way in which the instrument was constructed ensured its content validity, and item analyses on the instrument indicated an alpha reliability greater than .90.

The procedures used for the analysis of the data included three separate designs which were deemed appropriate to various aspects of the study. The three designs consisted of: 1). a four-facet factorial analysis of variance with repeated measures on one facet; 2).a three-facet factorial analysis of variance with repeated measures on one facet; and 3). a two-facet factorial analysis of variance with repeated measures on one facet. The final analysis involved the testing of fourteen a-priori hypotheses, relative to the above three designs, by means of orthogonal comparisons.

Of the fourteen a-priori hypotheses tested, two were found to be significant. One significant result indicated that working 15 or 30 exercises resulted in a greater gain in performance than working 5 exercises. The other significant result indicated that the students working 30 exercises with some feedback had a greater gain in performance than did the students working 30 exercises with no feedback. Two other hypotheses, though not significant at the .05 level, did indicate a trend in a given direction which resulted in the testing (on a post-hoc basis) of two additional hypotheses, both of which were found to be significant.

The primary conclusions to be drawn from this study are: 1). A larger number of exercises seems to result in a larger gain in performance than fewer exercises, as evidenced by the fact that the students working 15 or 30 exercises showed a significantly greater gain in performance than did those students working only 5 exercises. 2). The value of reinforcement may become greater as the number of exercises to be worked increases. This is concluded as a result of the finding which indicated that, at the 30 exercise level, those students getting some feedback gained more in performance than did those students getting no feedback; while no such difference was evidenced at the 5 or 15 exercises level. 3). If answers are to be used by the students, it is of no value to work the problems in their entirety. The results of this study seemed to indicate, very strongly, that detail in answers contributed nothing to performance gains.

BIBLIOGRAPHY

Bassler, O.C., Curry, D., Hall, W., Mealy, E. An Investigation of Two Instructional Variables in Learning Nonmetric Geometry. Professional Paper No. 69-10, Institute on School Learning and Individual Differences. Nashville, Tennessee: George Peabody College for Teachers, 1969.

Bourne, L. E., Jr., & Haygood, R. C. Effects of Intermittent Reinforcement of an Irrelevant Dimension and Task Complexity Upon Concept Identification. Journal of Experimental Psychology, 1960, 60, 371-375.

Bourne, L. E., Jr., & Pendleton, R. B. Concept Identification as a Function of Completeness and Probability of Information Feedback. Journal of Experimental Psychology, 1958, 56, 413-420.

Chansky, N.M. Learning: A Function of Schedule and Type of Feedback. Psychological Reports, 1960, 7, 362.

De Cecco, J. P. The Psychology of Learning and Instruction. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1968.

Gayne, R. M., & Staff, University of Maryland Mathematics Project. Some Factors in Learning Nonmetric Geometry. Monograph of Social Research in Child Development, 1965, 30, 42-49.

Gitman, D. A. Feedback, Prompting, and Overt Procedures in Nonbranching Computer Assisted Instructed Programs. Journal of Educational Research, 1967, 60, 423-426.

Hull, C. L. A Behavior System. New Haven: Yale University Press, 1952.

Koch, E. A., Jr. Homework in Arithmetic. Arithmetic Teacher, 1965, 12, 9-13.

Krumboltz, J. D. & Weisman, R. G. The Effect of Intermittent Confirmation in Programmed Instruction. Journal of Educational Psychology, 1962, 53, 250-253.

Maertens, N. An Analysis of the Effects of Arithmetic Homework Upon the Arithmetic Achievement of Third-Grade Pupils. Arithmetic Teacher, 1969, 15, 383-389.

Oppenheimer, D. B. The Relation Between Intelligence and Different Patterns of Feedback in a Linear Teaching Program. Journal of Experimental Education, 1968, 36, 82-85.

Skinner, B. F. Science and Human Behavior. New York: Macmillan, 1953.

Spence, K. W. Behavior Theory and Conditioning. New Haven: Yale University Press, 1956.

Walbesser, H. H., Jr. A Programmed Investigation of Fixed-Ratio Schedules of Reinforcement Using the System of Integers Under Addition. Unpublished Doctoral Dissertation, The University of Maryland, 1964.

A COMPARISON OF FOUR TYPES OF FEEDBACK TO STUDENT RESPONSES IN A CAI PROGRAM DESIGNED TO TEACH THE CONCEPT OF FUNCTION

Harold L. Schoen
Virginia Polytechnic Institute
and State University
Blackburg, Virginia 24061

In an attempt to help determine what types of CAI programs should be written for maximum instructional effectiveness, achievement and attitude scores of subjects receiving four different types of feedback in two CAI tutorial units were compared. Unit A consisted of set language, ordered pairs, and graphs. Unit B included the definition of function, graphs of functions, and functional notation. The units were written in I.B.M. Coursewriter III language and are available on one of The Ohio State University's I.B.M. 360/50 computer systems.

Two pilot studies were designed to determine the appropriate sample and to refine the instruments. In the final study sixty pre-calculuc mathematics students at The Ohio State University were randomly assigned in equal numbers to four cells, distinguished by the type of feedback the subjects received in Unit B, and were administered the CAI units and evaluation instruments via I.B.M. 2741 teletype terminals. The cells were the result of crossing two levels of each of two variables - I (individualization) and P (personalization). The levels of I were defined as follows.

I' - following an incorrect response, the student received feedback which gave the correct answer and stated why his answer was incorrect.
I''- following an incorrect response, the student received feedback which gave the correct answer and stated why it was correct, but did not refer to the student's incorrect answer specifically.

Feedback to correct responses was the same for the two levels of I. The levels of P were defined as follows.

P' - the student's first name appeared in some of the feedback (as often as was determined reasonable by the investigator).
P''- the student's first name never appeared in the feedback.

Thus the four treatment cells were I'P', I'P'', I''P', and I''P''.

Scores on two achievement tests on Unit B, O_1 and O_2, and an attitude test, O_3, all developed at least in part by the researcher, were the dependent variables. The Unit A treatment served to eliminate the "novelty of working on a machine" effect. A two-way analysis of variance was employed to analyze the data. The results showed no significant differences on O_1 scores, but the I'' (non-individualized) cells scored higher than I' cells on O_2 (p less than .04). The P' cells (those receiving their names in feedback) scored higher on the attitude scale than the P'' groups (p less than .07). In addition, on both O_1 and O_2 the P' cells scored higher, though not significantly higher, than the P'' cells.

The main implications of these results are:

1) In this study the least individualized feedback yielded highest achievement scores. Apparently more research is needed to determine better ways to achieve individualization of instruction via CAI if it is to be done.
2) Inclusion of a student's name in feedback to his responses in a CAI unit seems to improve both the achievement and the attitude scores.
3) To generalize beyond the mode of instruction, these results imply that the classroom teacher also should refer to students by name. Furthermore, contrary to common practice, there appears to be no advantage to commenting on a student's incorrect response in class or on homework. The teacher should simply tell the student he is wrong, state the correct answer, and explain why it is correct.

THE EFFECT OF COMPUTER PROGRAMMING EXPERIENCES ON STUDENT PROBLEM SOLVING BEHAVIORS IN EIGHTH GRADE MATHEMATICS

Thomas E. Foster
Sam Houston State University
Huntsville, Texas 77340

David C. Johnson
University of Minnesota
Minneapolis, Minnesota

This study reports data gathered to answer the question: Can treatments be constructed to incorporate student use of the computer and flow charts in the solution of non-standard problems, which will differentially affect performance on a series of test items that measure selected behaviors assumed to give evidence of problem solving ability?

A dependent and mutually reinforcing relationship is assumed between a student's mathematical maturity and his ability to solve non-standard problems. The importance then of showing the computer to be an aid in the development of the ability in students to solve non-standard problems, is more than specific, it is transitive in that the student's mathematical maturity would also be raised to a higher level.

Problem solving is a process which applies to the solution of non-standard problem settings. Such settings are non-routine in comparison to the drill and practice students encounter in much of their mathematical experience. A task is considered to be non-standard if the student has not previously demonstrated skill in producing acceptable solutions to it.

In solving problems one often makes use of selected supporting behaviors. The investigators identified nine such behaviors which are presumed to evidence the student's problem solving ability. These behaviors were:

(1) specifying conditions a datum satisfies
(2) selecting a relevant solution
(3) proposing a hypothesis
(4) identifying a pattern
(5) supplying missing information
(6) selecting relevant data (intermediate stage)
(7) using a constructed algorithm
(8) correcting an error in a constructed algorithm
(9) constructing an algorithm

These behaviors constitute, in part, the activity of programming a computer. It was hypothesized that a student who successfully programmed a computer to solve a series of non-standard problem tasks, would then further develop his problem solving ability, specifically in terms of the nine selected behaviors, in the process.

Four treatment groups within each of three intact eighth grade classes were provided with 24 non-standard problem tasks (12 required, 12 optional) over a 12-week period. These tasks were formulated to allow both computer and non-computer solutions. The treatments were in terms of the supplementary aids available: (1) neither computer nor flow charts, (2) flow charts only, (3) computer only, and (4) computer and flow charts. In addition to experimental groups, three base-line-data groups of eighth graders were defined; these students did not work on the non-standard problems or use the computer or flow charts. All students, experimental and base-line, were tested at the end of the 12-week period, using a 48-item test constructed to measure the nine behaviors described earlier. The base-line data groups were compared with the treatment group using neither the computer nor flow charts, to measure the effect the non-standard problem tasks had on the development of these behaviors measured by the post test. The treatment groups were then compared with each other, to ascertain the effect the use of computers or flow charts had on the development of these same behaviors. Three factors were present in the design: treatment, class, and reading level. Since the materials involved a reading component, students were blocked by high and low reading levels.

Since the base-line-data groups were not randomized populations, a descriptive statement only was made on their mean performance on the post test in comparison to that of Treatment Group 1 on this same test. Treatment Group 1 had a mean performance less than that of one base-line-data group and greater than that of the other two. On the sub-tests within the post test which measured the nine behaviors identified, Treatment Group 1 had a mean performance less than that of each of the three base-line-data groups on only two sub-tests and greater than that of each of them on only two sub-tests. Hence little difference was evidenced by this data and one might assume then that there was little or no effect due to the availability of the tasks alone.

A two-way analysis of variance on mean performance on the post test was used to ascertain the effects due to treatment, class, and reading level. Significant F-values were obtained on Reading within Treatment ($p = .01$) and Treatment X Class ($p = .05$). Pairwise comparisons using Scheffe's t-statistic were made, but this test failed ($p = .05$) to show that these significant F-values were within one treatment or reading cell within a treatment, though the high readers' mean performance was consistently higher than that of the low readers.

The mean performance of the four treatment (T) groups had the following directional order:

$$T1 < T2 < T4 < T3$$

Pairwise comparisons of the four treatment groups using a one-tailed

Dunnett's t-statistic were run to identify any significant difference between them. One such comparison yielded a significant t-value ($p = .05$); the mean performance of the group using the computer only was found to be significantly greater than that of the group using neither the computer nor flow charts.

A two-way multivariate analysis was run on mean performance on the nine sub-tests within the post test to find out which of these behaviors contributed to the significance or directions obtained in the two-way analysis of variance on the post test as a whole. The results of this analysis indicated that the significant differences found due to Reading within Treatment and Treatment X Class can be attributed primarily to mean performance on the subtests which measured the following three behaviors: (1) proposing a hypothesis, (2) identifying a pattern, and (3) selecting relevant data.

Two definite trends are suggested by the data obtained: (1) The computer is a useful aid in developing those behaviors presumed to evidence a student's problem solving ability. Each treatment group using the computer exhibited a greater mean performance on each of the nine behaviors separately than the group using neither the computer nor flow charts. On five of the nine behaviors the mean performance of the two computer groups was greater than that of the group using only flow charts. (2) Flow charts are also useful, in a more limited sense, as an aid in developing these same behaviors. The treatment group using only flow charts also had a greater mean performance on each of the nine behaviors separately than the group using neither the computer nor flow charts.

Tuesday, April 18, 1972

8:45 a.m. - 10:15 a.m. Bel Air Room

Research Reporting Section
NCTM Section No. 82

Subject: ASPECTS OF MATHEMATICS LEARNING

Presider: James M. Moser, University of Wisconsin--Madison

Speakers: Douglas A. Grouws, University of Missouri--Columbia
METHODS OF SOLUTION USED IN SOLVING OPEN SENTENCES

Lloyd F. Scott, University of California, Berkeley
CHILDREN'S NUMBER PREFERENCES: GEOMETRY VS. ABSTRACT

Schmuel Avital, The Ontario Institute for Studies in Education
LEARNING OF CONJUNCTIVE CONCEPTS IN MATHEMATICS IN THE JUNIOR HIGH SCHOOL

A. B. Durell, The Ontario Institute for Studies in Education
STRATEGIES FOR LEARNING MATHEMATICAL CONCEPTS

Reactor: Kenneth B. Henderson, University of Illinois--Urbana

METHODS OF SOLUTION USED IN SOLVING OPEN SENTENCES

Douglas A. Grouws
University of Missouri--Columbia
Columbia, Missouri

This investigation examined methods used by third-grade children in solving certain open sentences. The purposes of the study were 1) to identify the methods of solution employed by these children and to determine the frequency of use; 2) to ascertain the extent to which the same solution method was used by the children in solving different open sentences; and 3) to examine the relationship between number of correct solutions and the number and kinds of solution methods employed.

The study also examined certain relationships between methods used in solving open sentences and three factors (open sentence type, number size, context) associated with these open sentences. The open sentence type factor had four levels; each level reflected one of the following open sentence types: $N + a = b$, $a + N = b$, $a - N = b$, and $N - a = b$ where N was a placeholder for a number, and $\underline{a}$ and $\underline{b}$ were whole numbers. The two levels of the number size factor were related to the size of the whole numbers used as constants in the open sentences to be solved. The context factor was associated with the presence or absence of an appropriate verbal problem in an open sentence solving task.

Most of the reported research related to open sentences is in the classical framework of differences in difficulty within a set of basic facts. Very few studies have dealt with open sentences that were not in direct solution form (e.g., $N + 16 = 25$ and $16 - N = 9$). This study identifies methods of solution utilized by children when the open sentences presented are not in direct solution form. Thus, it provides important knowledge about a hereto unresearched facet of the ability to solve open sentences. Previous research (Weaver, 1970; Grouws, 1971) has shown that important differences in performance on open sentences of different types do exist, the present research provides information concerning methods of solution which appears to be related to these differences in difficulty.

Comparing methods of solution presented in classroom instruction with methods used by children in solving open sentences is essential if instructional techniques are to reflect how children learn. This study provides an initial step in this direction.

A sixteen item test was developed in which each test item consisted of an open sentence and in eight of these items a verbal problem was also included. The sixteen test items were completely crossed and balanced with respect to the three factors previously mentioned. The open sentence used in each test item was systematically generated. The whole numbers used as constants in these open sentences were randomly chosen from one of two number domains. The whole numbers in one domain were less than 19, and in the other domain the whole numbers were greater than 20 and less than 100. In eight of the test items an appropriate verbal problem was presented with the open sentence to be solved. In the remaining eight items only an open sentence was presented.

Sixteen boys and sixteen girls randomly selected from three Madison, Wisconsin elementary schools were individually given the sixteen item test. It was orally administered and each child was asked to solve the open sentence involved in each test item. Immediately after a child's numerical response he was questioned about the method he used in solving the open sentence. His numerical response was then recorded and his method of solution was classified according to an a priori developed eighteen category scheme. The classification scheme was developed by first listing methods children might use in solving open sentences. Following this logical analysis the scheme was pilot tested, expanded, and then revised to arrive at the final form used in the testing.

Two weeks after the completion of the interviewing, coding reliability was checked. The interviewer recoded 50 selected items by listening to audio recordings of the appropriate parts of the original interviews. This second classification of solution methods was in 92 percent agreement with the original coding. That is, in 46 of the 50 tasks considered the associated solution methods were classified in the same category as they had been originally.

The data generated from this study were more qualitative in nature than quantitative and were, therefore, handled primarily through the use of descriptive statistics.

Solution methods ranging from the relatively intuitive (e.g., tallying and counting) to the relatively formal (e.g., use of an operation and use of an inverse relationship between operations) were identified. The four most frequently used methods were: Operation (196 times), recall (67), counting (57), and substitution (38).

The mean number of different solution methods used on the sixteen item test was 5.12. The variance was 2.56 and the range was from 1 to 9 different methods. Students who performed well (15 or 16 correct responses) used an average of 2.75 different methods. In 49 or 77.8 percent of the problems situations presented to these children a solution method classified as use of a formal operation was used. The number of different solution methods used and the number of correct responses given were inversely related and moderately correlated (i.e., coefficient of -.48).

Formal addition or subtraction (i.e., operation classification) was used 136 times in connection with two-digit items and 60 times in connection with basic fact items. Recall was not used on the two-digit items but was used 67 times on the basic fact type items.

(1) Children (at this level) use a wide variety of methods in solving open sentences of the types considered.

(2) There is considerable variation in the number of different solution methods used by different children in attempting to solve the same set of open sentences.

(3) Children's solving performance and the number and kind of solution methods employed are substantially related.

(4) Solution methods used by children are often not the methods presented in classroom instruction.

BIBLIOGRAPHY

Grouws, Douglas A. Some Factors Associated With Children's Solving Performance on Four Types of Mathematical Open Sentences. Paper presented at the 1972 Annual Meeting of the American Educational Research Association.

Weaver, J. Fred. Some Factors Associated with Pupil's Performance Levels on Simple Open Addition and Subtraction Sentences. _Arithmetic Teacher_, 1971, 18, 513-519.

19

CHILDREN'S NUMBER PREFERENCES: GEOMETRY VS. ABSTRACT

Lloyd F. Scott
University of California
Berkeley, California 94720

The investigation was undertaken to determine the relative stability of young children's use of abstract and concrete number representations. An analysis of currently available curriculum materials for grades K through 6 reveals what may be regarded as an obsession toward early abstraction, symbolization and manipulation of symbols. In most cases, written numerals are introduced in the kindergarten program and/or within the first few lessons of first-grade programs, and more complex symbolic representations follow soon thereafter. This practice has been maintained over the recent history of school mathematics program reform despite the increasing number of cautions from learning theorists and researchers in mathematics education concerning premature abstraction. Quite apart from speculation as the effect of this practice upon attitudes toward mathematics, upon a view of the nature of mathematics or upon mathematical achievement, its relationship to the nature of children's thinking should be explored. Are these symbols stable representations for mathematical concepts, or are they, as suggested by Piaget's work, unstable artifacts of abstract experience until the concrete experience has matured? Do children as they get older demonstrate greater reliance upon more abstract representations of number ideas, or does their trust remain with more concrete representations? At what point in children's usual experience in school mathematics may the teacher assume that the concrete experience has matured and the more abstract mathematical representations are stable?

The investigation sought answers to these questions through a test of the following hypothesis:

With respect to the natural numbers, children between the ages of 5 years and 12 years will demonstrate an increasing reliance upon and preference for more abstract mathematical representations as they get older.

The total investigation proceeded through several phases extending over four years and involved a total of 2,411 children, grades K through 6, all in elementary schools in Berkeley, California. Only the final phases of the study involving improved procedures and 1,843 children shall be reported here.

It was learned early that the measure of children's preferences should involve if possible a non-verbal, behavioral task. The procedure finally chosen was one of sorting cards as follows:

The child is given three cards to sort according to models developed by the interviewer and left exposed for the child's reference. It was established that the child was to follow the pattern of the demonstrated models, although no verbal instruction on the nature of the pattern was given. It was expected that the child would observe a less than or smaller than arrangement toward the right (see Figure 1), but whatever the child perceived as the pattern was allowed to stand unchallenged. The models and their arrangement took the form shown in Figure 1, although the order was changed in various parts of the total investigation to determine the influence of mental set and to determine the reliability of the technique.

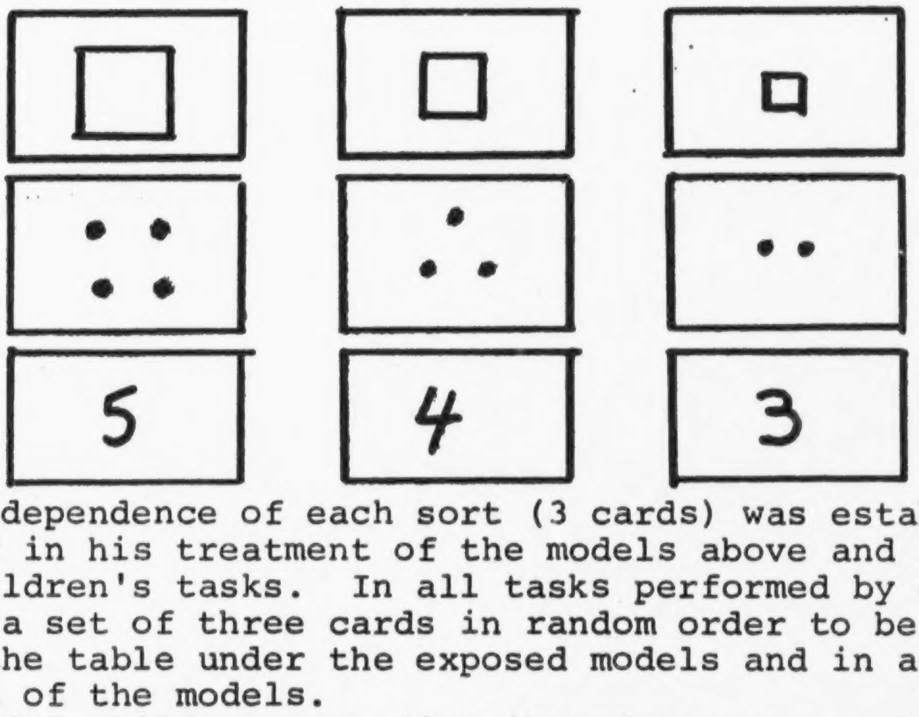

Figure 1

The independence of each sort (3 cards) was established by the interviewer in his treatment of the models above and in his treatment of children's tasks. In all tasks performed by children, each was handed a set of three cards in random order to be arranged and placed on the table under the exposed models and in accord with the pattern of the models.

The 1,843 children were distributed rather regularly through the seven grade levels, kindergarten through grade 6, and data summaries were made for each grade level. The summaries in the report will deal with the children's responses to eleven separate tasks, four of which will be treated in detail. Six of the tasks were designed to reinforce the models and offer information as to the child's perception of the pattern. All other tasks were designed to put the child into a condition of forced choice between a relatively concrete attribute and a relatively abstract attribute. One of these tasks was a fairly direct Piagetian task as follows:

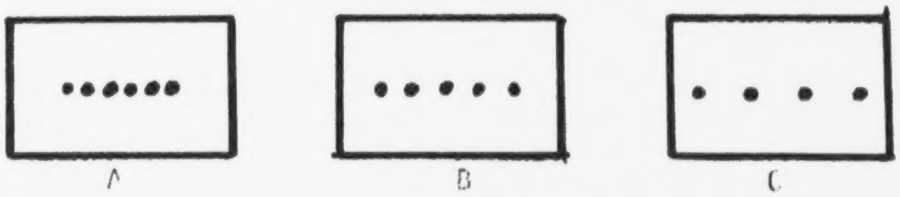

Figure 2

It may be seen that if the child opted for the A B C arrangement as shown in Figure 2, the number of discrete dots may be regarded as the operational stimulus. If he chose the C B A order, he may be assumed to be responding to the length of the row, a more concrete characteristic than that of abstract number.

The other four tasks of special interest were similarly conceived.

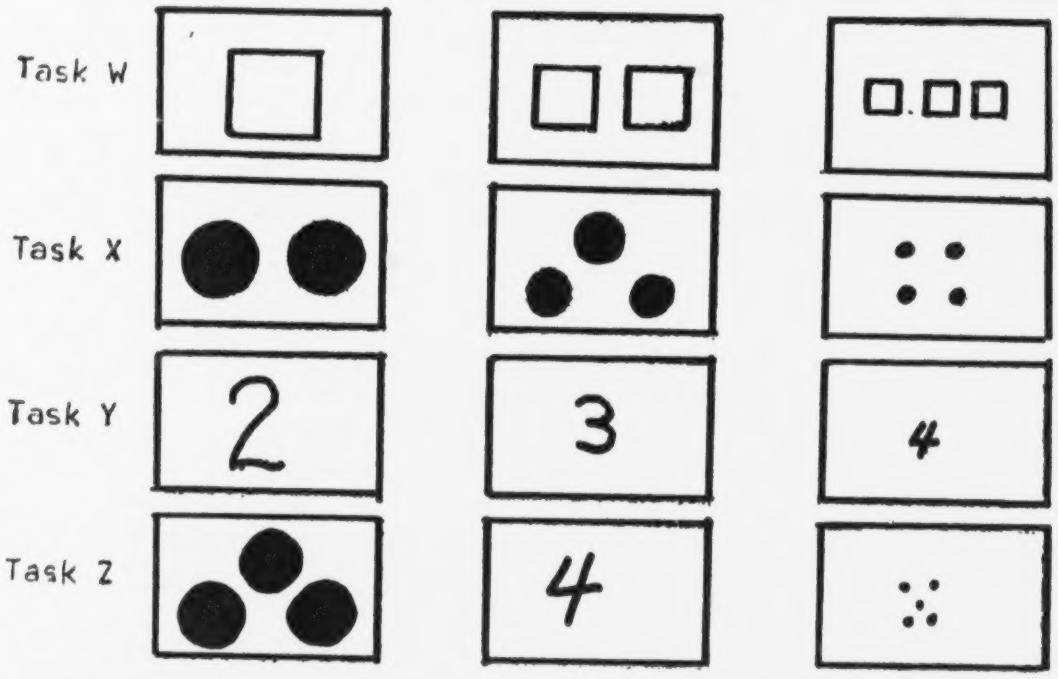

Figure 3

It will be noted that in the order shown for each of the tasks W - Z in Figure 3, the predominant stimulus would appear to be the more concrete of the choices. If the order in all cases were exactly reversed, the more abstract stimulus would seem to prevail.

Post experimental interviews were held with a number of the children to expose non obvious ordering criteria for responses otherwise categorized as random responses and to shed further light upon the nature of children's thinking about the natural numbers.

The results of the investigation offer some interesting food for thought. While it had been hypothesized that there would be an increasing acceptance of (preference for) the more abstract representations of number as children became older, this characteristic was not revealed in the result. The preference for the concrete seems to decrease only very gradually as age increases during the range studied here, and a post hoc extension in a graduate teacher training program reveals that even students much older will opt for the more concrete stimulus in a forced choice situation. While the job of interpretation and drawing summary conclusions is not yet complete, it would seem that there may be some minor conflict between this evidence and the popular conception of Piaget's age normative stages of cognitive development. Undeniably there is some striking conflict between this evidence and some prevailing practices in the selection of mathematics textbook activities for young children.

22

LEARNING OF CONJUNCTIVE CONCEPTS IN MATHEMATICS IN THE JUNIOR HIGH SCHOOL

Schmuel Avital
The Ontario Institute for Studies
in Education
Toronto 181, Ontario, Canada

and

(Miss) Malca Grossman
Israel Institute of Technology

The study is concerned with the thinking process of school children in learning mathematical concepts within the scope of their knowledge.

The objects of the study were:

(a) Detection of strategies in learning mathematical concepts.
(b) Verification of the psychological status of these strategies, and the degree of consistency in their application.
(c) Detection of a relation between the modes of learning on the one hand, and the cognitive variables on the other.
(d) Examination of reflections in terms of the teaching process, and ways for its improvement in the light of the results.

The participants were 12 girls and 12 boys--seventh-grade students, age group 12-13--drawn from the upper 40% of the class.

In three experiments, the participants were given an array of instances--some of them positive, some negative, as explained below--representing all possible combinations of 4-5 attributes with two values (positive and negative) each, and required to identify a conjunctive concept represented by a combination of two attributes, using the most effective technique possible. The identification process began with a positive instance and consisted in successive choices from the array, each choice being characterised according as it did or did not represent the concept.

The first experiment involved a 32-number array constructed according to the following attributes: evenness, squareness, divisibility by 3, value of digit relative to 5, value of number relative to 200. Four problems were to be solved, two in the full-scale array and two in a smaller one comprising 16 numbers, all of them smaller than 200. Half the group began with larger array and half with the smaller one.

The techniques used by the participants were rated against four ideal strategies: conservative focusing, focus gambling, simultaneous scanning, and successive scanning--those of the focusing type being based on examination of the attributes, and those of the scanning type--on hypotheses regarding the nature of the concepts.

The principal findings of the experiment were as follows:

17 participants adopted a focusing strategy and 7 a scanning strategy. The second category was superior in its average I.Q. and also (although not significantly so) in its average marks in mathematics and language. In the case of the smaller array, the scanners showed superiority in terms of average number of solved problems; in that of the larger one, the focusers showed superiority despite their "inferiority" from the cognitive ability viewpoint.

From the viewpoint of performance--(measured as the solution time) the number of instances required, and the number of redundant choices--the focusing strategies were superior.

The relation between performance and cognitive ability was studied by several means. Correlation coefficients indicate a correlation between performance and I.Q. in several cases. At a significance level of 0.05, data did not justify rejection of the hypothesis denying a correlation between performance on the one hand and achievement in mathematics and language on the other.

Several characteristic tendencies were detected in the acquisition and processing of the information: reluctance to use indirectly-acquired data, an influence of the participant's expectation on his technique, etc. Characteristic mistakes were also noted.

The participants who began with the smaller array showed gradual improvement in its successive problems, no such improvement was shown by those who began with the larger array.

The second experiment (20 participants in 4 subgroups) involved 16 tables of operational systems representing all possible combinations of the following attributes: commutativity, associativity, and identity element, and a complement-to-1*. Results were as follows: in the first individual stage 75% of the problems were solved, as compared with the 16-number array--indicating progress between the experiments. The focusers showed

---

In the (S,) system the complement-to-1 of an element x of a set S is an element y of this set satisfying x*y=y*x=1. In systems where 1 is the neutral element, the "complement-to-1" is simply the inverse.

superiority in the number of solved problems, and the scanners--in performance (solution time and number of choices).

No significant relation was found between performance and the cognitive ability variables.

In the teamwork stage all subgroups attained a correct solution with considerably superior performance (mainly from the time viewpoint) compared with the average of its members as individuals. The four subgroups decided on the conservative-focusing strategy after a consultation within each.

In the additional individual stage, undertaken by two of the subgroups, focusing was resorted to throughout. Improved performance was observed compared with the first individual stage.

The third experiment (10 participants) involved an array of 32 fractions classified according to the following attributes: positive-negative; reduced-unreduced; denominator 10; proper-improper; periodic-terminating. The assignment consisted in identifying the concept "proper periodic fraction." No clue was given as to the underlying principle of the array, but the participants were allowed to inquire whether or not a certain attribute was relevant, and encouraged to consult among themselves and decide jointly on the choices. They eventually identified the five relevant attributes (periodicity was identified with the aid of a cue on the various forms of representation of numbers), and in due course arrived at the concept in a systematic and efficient manner without logical errors, based on conservative focusing.

The study indicated that the learning process differs from student to student, a fact which dictates diversification of the class technique and its adaptation to these individual requirements. The study also brought out the need for training in mathematical methodology and sound logical thinking, in addition to presentation of the subject matter.

The strategy adopted was associated with an individual style of thinking. Consistency in choosing a strategy was observed in the majority of the cases. At the same, the feasibility was demonstrated of training in strategies, thereby improving the students' learning technique.

A new form of concept teaching--via instances--was indicated, and its advantages were proved in promoting mathematical thinking, improving insight, and providing motivation and interesting drill.

The study also indicated achievements in self-supervised independent teamwork, free from intervention by the teacher.

In conclusion, the present exploratory study yielded a paradigm for controlled experimental investigation of the cognitive processes in learning significant concepts.

STRATEGIES FOR LEARNING MATHEMATICAL CONCEPTS

A. B. Durell
The Ontario Institute for Studies in Education
Toronto 181, Ontario, Canada

A study was conducted to determine if instruction in the use of a concept learning strategy would affect the strategy use or learning efficiency of mature subjects.

Studies in concept learning have revealed two basic types of strategies used by learners. One is called focussing. A learner who uses a focussing strategy will use a given positive instance of the concept as a basis for his investigation. He will vary one or two of the attributes of this focus instance on each trial and thereby accumulate information about the relevance of each attribute.

The other type of strategy is called scanning. A learner using this strategy will test complete hypotheses until he has logically eliminated all but one.

Studies of concept learning, which take strategy use into account, usually seek to discover which strategies subjects choose to use and how effectively they use them. In this case subjects were taught either a focussing or a scanning strategy or were assigned to a control group which received no strategy training.

The determination of what strategy was used has commonly been a subjective judgment. It is possible to score the number of differences of the attribute value of successive instances from the attribute values of the given focus instance. The fewer the number of differences, the more the subject is focussing. Thus the mean of the number of differences gives an objective measure of the degree of focussing. At the same time, the variance of the number of differences is a measure of the consistency of strategy use. A computer program was used to perform the tedious task of counting the differences.

In most concept learning experiments, it is common for the subject to verbalize the concept for the experimenter when he thinks that he knows it. Given the present state-of-the-art in computer processing of natural language, such a procedure should be avoided in computerized experiments. In this case a pseudo-mathematical task and a means of indicating knowledge of the concept were developed which would exploit advantages of computers and also avoid their limitations.

A typical stimulus is (3,4,1):3. The digit following the colon is the "result." It is produced from precisely two of the three digits within the parentheses by addition, subtraction, or multiplication. The two digits used to produce the result may be selected according to their position within the triple of digits, or according to their relative magnitude, or by a combination of the two. Instead of selecting the next instance from an array or being presented with one of the experimenter's choosing, the subject constructs his own, types it in, and is informed whether it is a positive or a negative instance of the concept. When he thinks that he knows the concept, the subject requests a test. The test presents him with triples of digits for which he must produce the correct result. When he answers five consecutive test items correctly, it is assumed that he has the concept.

Subjects were 60 graduate students in education. They were assigned in groups of 20 to one of the experimental groups.

Each subject worked seated at a teletype terminal. Each received the same introductory information. This was followed by instruction in the strategy to be used or, in the case of control subjects, practice in finding two sample concepts. Typed-in responses were recorded automatically for analysis after completion of the experiment.

Analysis of variance indicated significant differences among the mean degree of focussing scores, and among the mean scores for consistency of strategy use of the treatment groups. No significant differences among the mean number of trials to criterion were found.

The degree of focussing and consistency of strategy use data were examined further using the Tukey method for multiple comparisons. This analysis indicated that instruction in the use of a strategy did affect the degree of focussing which subjects displayed. Only subjects taught the focussing strategy showed more consistent strategy use than the untrained subjects did.

The study produced more questions than answers about concept learning but the task developed, and the program for objective determination of the degree of focussing and consistency of strategy use should be potent tools for further investigation of concept learning. The task is a compromise between the abstraction of traditional laboratory concept learning tasks and the familiarity of classroom mathematical concepts. Elimination of the need for verbalization makes the experiment amenable to computer administration.

Tuesday, April 18, 1972

12:15 p.m. - 1:45 p.m. Bel Air Room

Research Reporting Section
NCTM Section No. 114

Subject: GENERAL RESEARCH TOPICS IN MATHEMATICS EDUCATION

Presider: Kenneth B. Henderson, University of Illinois--Urbana

Speakers: Alan W. Holz, Purdue University
A TECHNIQUE FOR STUDYING THE ORGANIZATION OF MATHEMATICS TEXTS

Gerald Kulm, Purdue University
NEGATIONS IN LINEAR SYLLOGISMS AND PROBLEM SOLVING

Constance Anne Carroll, Edgecliff College
LOW ACHIEVERS' UNDERSTANDING OF FOUR LOGICAL INFERENCE FORMS: AN ANALYSIS OF DIFFICULTIES AND OF THE EFFECT OF INSTRUCTION

Susan Rudder Fuller, Flint River Academy
A STUDY OF PROBLEM SOLVING METHODS OF STUDENTS WITH AND WITHOUT THE CONSTRAINT OF A TIME LIMIT

Reactor: Richard J. Shumway, Ohio State University

A TECHNIQUE FOR STUDYING THE ORGANIZATION OF MATHEMATICS TEXT

Alan W. Holz
Purdue University
Lafayette, Indiana 47907

Comparative studies of mathematics instructional materials in which the criteria were learning differences have suffered from lack of well defined treatment differences, from expanded development effort spent on experimental treatments, and from the presence of confounding variables such as teacher influence and between-subject interactions. Other research designed to systematically study mathematics text materials is sparse. Only in a few areas such as readability and sequencing has systematically study mathematics text materials is sparse. Only in a few areas such as readability and sequencing has systematic research related specifically to textual materials been carried out. Additional significant variables in mathematics text materials need to be identified and methodologies for investigating these variables need to be developed.

This paper concerns the development of a technique for identifying and studying organizational variables in mathematics text in terms of the semantic structure of the materials. The structure of any object may be defined by its elements and by the interrelationships between these elements. These interrelationships concern the sequencing or ordering of the elements and the frequency with which the various elements occur. When considering the structure of mathematics text the elements of the structure may be thought of as *messages*. The interrelationships between the messages determine the *organization* of the material.

A mathematics text passage can be thought of as consisting of a series of messages. In terms of formal mathematical content, the basic messages can be conceived of as undefined terms, definitions, axioms, and theorems. Definitions require examples and theorems require explanations which are called proofs. For the mathematics to be meaningful applications are required. To help the student learn the mathematics, exercises and problems are usually provided. In terms of presentation of the content, the basic messages can be conceived of as consisting of ordinary words, symbols, or various types of illustrations. These considerations of mathematical content and mode of presentation have led the author to develop a two dimensional category system for classifying messages in mathematics text. This system of categories is given below.

Classification of Messages in Mathematics Text (CMMT Categories)

Dimension 1: Content Analysis

	0. Blank Space	
ssages quiring ly ception.	1. Definition (Meanings of words or symbols.) 2. Generalizations (Important rules, axioms, theorems, formulas, etc.) 3. Specific Explanation (Concrete examples and discussion.) 4. General Explanation (Proofs of theorems, general discussion, etc.)	Exposition
sages ling ponses er n eption.	5. Procedural Instructions (Directions.) 6. Developing Content (Questions in exposition, developmental activities, etc.) 7. Understanding Developed Content (Exercises requiring practice, computation, identification, etc.) 8. Applying Developed Content (Real world problems, using generalizations in concrete situations, etc.) 9. Analyzing and Synthesizing Developed Content (Proving propositions, finding new relationships in developed content, etc.)	Subclassifications of questions, exercises, problems, activities, etc.
	10. Other Material (Headings, non-mathematical materials, etc.)	

Dimension 2: Presentation Mode

0. Blank Space

Written text.

1. Words
2. Mathematical Symbols

Illustration

3. Representations of Abstract Ideas (Venn diagrams, geometric diagrams, mapping pictures, etc.)
4. Graphs (Number lines, coordinate graphs, bar graphs, etc.)
5. Representations of Physical Objects of Situations (Plans, maps, cross sectional drawings, photographs, etc.)
6. Non-mathematical Illustrations (Motivational cartoons of photographs, etc.

7. Combinations of Illustrations with Written Text (Flow charts, mathematical tables, tree diagrams, etc.)

The CMMT category system is applied to mathematics text passages in much the same way that many other observation scales are applied to the phenomenon they measure. The basic unit of measure used is one-third of a line of print. An area unit is used so that quantitative aspects of the organization of passages can be described in terms of the space devoted to the various types of messages. To apply the category system to a page of text a grid conforming to the format of the page is drawn. This grid partitions the page into units and a category number for each dimension of the CMMT system is recorded for each unit.

When classification of all messages in a passage is completed the information may be analyzed in ways which reflect both sequential and quantitative aspects of the organization of the passage. Sequential aspects are represented by making an ordered list of the classifications following the natural flow of the printed material through the grid. A matrix similar to those used in Flander's interaction analysis is used to analyze the nature of the interactions among CMMT categories. Quantitative aspects of a passage are described by determining the proportion of messages in various categories and in logical combinations of categories. All of the descriptive information for a given passage is called the CMMT analysis for that passage. A computer program is used to derive the CMMT analysis of a passage from a rater's sequence. Examples of CMMT analyses of selected passages are included in the paper.

Thorough descriptions of the CMMT categories have been made and specific rules for raters using the CMMT category system to rate passages have been developed. These rules describe general procedures for raters to follow and give specific decision rules for dealing with such problems as the classification of units containing more than one type of message and what to do when in doubt about the classification of a message. A method which has been used to estimate reliability of raters using Flander's interaction analysis has been adapted for estimating the reliability of CMMT raters. Estimates of rater reliability are given in the paper.

The CMMT technique described here has impact for both practical application and research. CMMT analysis provides a means of describing the organization of existing mathematics text materials in quantitative terms. It provides information which could be of value to textbook selection committees, publishers and authors of mathematics texts. CMMT analysis provides a method for manipulating and controlling organizational variables in comparative studies of written mathematics materials. Thus it gives promise of becoming an effective instrument for studying mathematics text materials.

NEGATIONS IN LINEAR SYLLOGISMS AND PROBLEM SOLVING

Gerald Kulm
Purdue University
Lafayette, Indiana 47907

The processing of negations has been of considerable interest to psycholinguists for some time. In particular, there has been recent research and resulting controversy about negations in linear syllogisms. For example, how does a person process: John is older than Joe, Bill is not as old as Joe. Who is oldest?

In mathematics, linear syllogisms often occur in problems such as: John is 3 years older than Joe, Bill is 2 years younger than Joe. If the sum of the three boys' ages is 22, how old is John? The student must solve the syllogism in order to solve the problem. Previous investigations of syllogistic reasoning have not considered the effect of a numerical problem on solving a related syllogism. Researchers in problem solving have not considered the effect of different forms of syllogisms on solving the related problem. The purpose of the present study was to explore the effect of negations in linear syllogisms on solving numerical problems that are based on these syllogisms.

Linear syllogisms can be represented as follows where p, q, and r are variables and R is some relation.

(1) pRq, qRr. Which is 'R'est?
(2) pRq, q notR r. Which is 'R'est?
(3) p notR q, qRr. Which is 'R'est?
(4) p notR q, q notR r. Which is 'R'est?

The question "Which is not 'R'est?" can also be asked, giving eight possible syllogisms, not counting the permutations of p, q, and r.

Researchers generally have used latency of response as the criterion and have found that (1) takes the least time to solve, (4) takes longest, and (2) and (3) are somewhere between. It was hypothesized in this study that similar latency measures would be obtained. After a subject has solved a negated syllogism, does he use the solution in solving a related numerical problem? Does the form in which the syllogism is stated affect the strategy that the subject uses in solving the problem? Does the subject solve the syllogism over again each time he refers back to the given information in the problem? The answer to the first two questions can be inferred from an investigation of the kinds and orders of the arithmetic operations and the errors made during problem solution. The answer to the last question can be inferred from the total time taken to solve the problem.

Subjects: The subjects were 40 eighth-graders from a suburban New Jersey school. Testing was done individually by the experimenter.

Materials: The following eight problems were used.

Treatment I (1) p is smaller than q, q is smaller than r. Find p.
(2) p is smaller than q, q is smaller than r. Find r.

Treatment II (3) p is not as large as q, q is smaller than r. Find p.
(4) p is not as large as q, q is smaller than r. Find r.

Treatment III (5) p is smaller than q, q is not as large as r. Find p.
(6) p is smaller than q, q is not as large as r. Find r.

Treatment IV (7) p is not as large as q, q is not as large as r. Find p.
(8) p is not as large as q, q is not as large as r. Find r.

The capital letters A, B, and C were substituted randomly for p, q, and r in each problem and the problems were then typed on 3 x 5 cards.

On a piece of cardboard, the following expressions were printed: A, B, C, A + B, A - B, B - A, A + C, A - C, C - A, B + C, B - C, C - B, A + B + C. Opposite each expression was a "window" that could be opened by pulling a tab. A set of numerical values were prepared for each of the eight problems to match each of the above expressions, except for one of A, B, or C which was the unknown.

After doing a practice problem, S was given a card from one of the pairs I, II, III, or IV. Within the pairs, the order was counter-balanced so that 5 Ss did 1 then 2, 5 Ss did 2 then 1, and so on for each pair. The latency between receiving the problem card and pulling the first tab was timed to the nearest tenth of a second with a stopwatch. As S proceeded with the problem, the order in which the windows were opened was recorded. As soon as S gave the correct number for the unknown, the total time was recorded. If S did not solve the problem, four minutes was entered as the total time. The procedure was then repeated for the second problem.

Seven variables were selected to represent the strategy used by the Ss. These will be called process variables and are described as follows: (1) looked at a number first or a relation first, (2) subtracted incorrectly, (3) used subtraction rather than addition to solve, (4) solved as soon as sufficient information had been obtained, (5) made an incorrect guess, (6) used the first obtained information in the solution, and (7) used one of p and q, q and r, or p and r in the solution.

Three outcome variables were used: (1) latency of response, (2) total solution time, and (3) number of steps, i.e., number of "windows" opened.

Each of the ten variables was entered into a 4 x 2 x 2 (Treatment × Order × Directions) ANOVA, where directions (Find p or Find r) was a within Ss variable.

Outcomes: For response, the treatment effect was significant ($F_{3,32} = 2.93$, $p < .05$). Planned comparisons tests revealed that the latency for Treatment I was significantly ($p < .01$) smaller than for treatments II, III, and IV which did not differ significantly. None of the other effects were significant for latency. For total solution time, there were no significant effects and for number of steps, only the order × directions interaction was significant ($F_{1,32} = 4.75$, $p < .05$).

Processes: Of the process variables, significant differences were obtained only for variable (7), use of p and q, q and r, or p and r in the solution. A significant main effect was found for directions ($F_{1,32} = 16.33$, $p < .01$) and for treatment × order × directions interaction ($F_{3,32} = 4.31$, $p < .05$).

The results indicated that response latency was similar to results of previous studies. Since there were no differences in total solution time for the treatments, it may be inferred that subjects transformed the negated statements and used the transformed statement in solving the problem. Further evidence for this transformation was the lack of significant differences between treatments for the process variables.

The interactions of problem order and directions for the "number of steps" and "terms used" variables provided some evidence that finding the largest term facilitated finding the smallest but not vice versa.

This study showd that it is possible to apply results from studies in psycholinguistics to the language and symbolism of mathematics. It is necessary, however, to use caution in generalizing these results to the special difficulties involved in comprehending the language of mathematics in problem solving settings.

LOW ACHIEVERS' UNDERSTANDING OF FOUR LOGICAL INFERENCE FORMS: AN ANALYSIS OF DIFFICULTIES AND OF THE EFFECT OF INSTRUCTION

Sister Constance Anne Carroll
Edgecliff College
Cincinnati, Ohio 45206

The ability to draw logical inferences is a necessary, although not sufficient, condition for success in mathematics. As students strive to master the techniques of understanding and formulating mathematical arguments, it is essential that they be able to distinguish valid deductions from invalid ones.

Recognition of the bond between logic and mathematics is observable in theory propounded by Jean Piaget (1968), who holds that "all mathematical reasoning is formal, or as the logicians put it, hypothetical-deductive." This formal deduction, Piaget explains, "consists in drawing conclusions, not from a fact given in immediate observation, nor from a judgment which one holds to be true without any qualifications...but in a judgment which one simply assumes...just to see what it will lead to [p. 69.]" Far from being a form of thought of little use to children, this type of reasoning is constantly required in solving simple problems, often found in mathematics texts, which begin with a phrase such as "If one orange costs ten cents...." In dealing with such problems, the child is forced, according to Piaget to

> reason in conformity with premises that are simply given, which means disregarding reality and even suppressing any memories or real observations which might block the way to the process of reasoning. Reasoning of this kind is done from pure hypotheses... Either a mathematical problem is presented to a child as a purely empirical problem, in which case he will be kept in ignorance of the deductive power of even elementary arithmetic or he will be compelled to reason strictly, but then this process, in so far as it appeals to fixed rules and to previously admitted propositions, will be a process of formal reasoning [p. 69].

Elsewhere, Piaget and Inhelder (1958) assert that

> ...the role of possibility is indispensable to hypothetico-deductive or formal thinking.... The connection indicated by the words 'If...then' (inferential implication) links a required logical

> consequent to an assertion whose truth is merely a possibility [p. 257].

It is precisely this form of hypothetical reasoning from given premises, in particular from conditional sentences of the form "If...then" that is the focus of the present study. Such sentences are common in mathematics, yet the precise meaning and implications of such sentences are often obscure to the pupil. The present study is concerned with two issues related to conditional reasoning. The investigator sought:

1. To examine the effectiveness of instructional sessions with small groups as a means of improving performance in conditional reasoning;
2. To analyze the difficulties that low achievers experience in conditional reasoning.

Four forms of conditional reasoning paradigms, the modus ponens, converse, inverse, and contrapositive, were considered. Four content dimensions, familiar, symbolic, misleading, and removed from reality, were used in each of these forms.

Subjects for the study were ninth-grade students classified as low achievers in mathematics. Six groups of subjects were formed: a girls' and boys' experimental group, a girls' and boys' alternate treatment group, and a girls' and boys' contrast group. Both the alternate treatment and the contrast groups served as control groups.

A conditional reasoning test, constructed by the experimenter, was administered to all subjects before and after treatment. The subjects in the experimental groups received instruction consisting of six lessons in conditional reasoning; the subjects in the alternate treatment groups received instruction consisting of six lessons in probability; subjects in the contrast groups received no instruction. After all subjects had taken the posttest in conditional reasoning, a test measuring a reflective-impulsive dimension of cognitive style was administered to the subjects in the experimental groups. (The test used was an adaptation by Kagan of his "Matching Familiar Figures" test.) Data from the above measures were used to examine the hypotheses and questions proposed in the study. The following findings emerged.

1. The improvement among subjects in the experimental groups, measured by the percent of the subjects whose total score was higher on the posttest than on the pretest, was not significantly greater than that among students in the control groups ($\alpha = .01$).
2. The percent of subjects in the experimental groups who improved on the converse form of argument was significantly higher ($\alpha = .01$) than in the control groups. Differences among the percents showing improvement were not significant for the other three forms, but the trend favored the experimental groups on the modus ponens and inverse, and the control groups on the contrapositive.

3. The percent of subjects in the experimental groups who improved on the misleading and removed from reality content dimensions was significantly higher ($\alpha = .01$) than in the control groups. Differences among the percents showing improvement were not significant for the familiar and symbolic dimensions, but the trend favored the experimental groups.
4. Subjects who had not had instruction in conditional reasoning made significantly more errors on the fallacious arguments (converse and inverse) than on the valid arguments (modus ponens and contrapositive).
5. Test results indicated that for subjects who had no instruction in conditional reasoning the modus ponens was the easiest form, followed by the contrapositive, with little distinction between the converse and the inverse, which were the most difficult forms. (These conclusions were based on average scores and on percent of subjects having K or more items correct on each category according to form. No significance tests were made regarding the results reported in "5".)
6. The effect of content seemed to vary among the different forms. On the modus ponens form the order of increasing difficulty of the content was familiar, removed from reality, symbolic, and misleading. On the coverse and inverse forms the order was familiar, misleading, symbolic, and removed from reality. On the contrapositive form the order was removed from reality, symbolic, familiar, and misleading. (The order was determined by the percent of subjects having K or more items correct in each category according to form and content combination. No significance tests were made.)
7. Correlations between scores on the conditional reasoning and cognitive style tests were all very low, giving little evidence of any linear relationship between reflection-impulsivity and conditional reasoning skills.

Thus it can be seen that the factors influencing the subjects' attempts to reason logically are many. Knowledge concerning any of these factors could shed further light on the reasoning process. This process, important to all engaged in education, is especially pertinent to the task of the mathematics educator, who must be involved in helping students to develop skill in drawing conclusions, in forming and recognizing proofs, and in avoiding fallacious reasoning.

REFERENCES

(Only those works referred to in the summary are here included; the complete list of references for the study is quite extensive.)

Inhelder, B., & Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958.

Kagan, J., Rosman, B.L., Day, D., Albert, J. & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. Psychological Monographs, 1964, 78 (1, Whole No. 578).

Piaget, J., Judgment and reasoning in the child. Totowa, N.J., Littlefield, Adams, 1968.

A STUDY OF PROBLEM SOLVING METHODS OF STUDENTS WITH AND WITHOUT THE CONSTRAINT OF A TIME LIMIT

Susan Rudder Fuller
Flint River Academy/Woodbury, Georgia 30293

The purpose of this study was to determine if there are students who use different methods in solving mathematics problems under the constraint of a time limit as compared to the methods they employ when they are not under such constraint; and if such students exist, to determine how they differ on three selected variables from students who use the same methods when solving problems both with and without a time limit.

Sixty-four subjects of average and above-average intelligence were individually administered two problem solving tests. On one test the subjects were told that they had three minutes in which to solve each problem, and on the other test the subjects were told that they could have as much time as they needed to solve each problem. Subjects were asked to think aloud as they solved each problem, and the problem solving interviews were recorded on audio tape.

The problem solving processes each subject employed were analyzed and coded by the investigator from the recordings of the interviews. In the coding system employed, problem solving processes were grouped into four categories: (1) reading the problem (r), (2) re-reading the problem (R), (3) deduction (D), and (4) trial and error (T).

In order to determine if a subject changed problem solving methods underthe two time conditions, a pattern of problem solving was computed for each subject on both the timed test and on the untimed test. A subject's pattern of problem solving on a test consisted of four probability estimates of being in each of the four problem solving processes identified by the coding system. A subject's pattern of problem solving on the untimed test was compared to his pattern on the timed test by a contingency table and the chi square statistic. A subject was designated as changing problem solving methods on the timed and untimed test if his two patterns of problem solving differed at the .01 level of significance.

Four hypotheses were tested at the .05 level of significance, and the following conclusions were reached.

The first hypothesis was that the number of subjects who change problem solving patterns was not significantly different from zero. Twenty-one subjects were identified as changing problem solving methods under the two time conditions, and it was concluded that this was a significant number of subjects who changed methods on the two problem solving tests under the two time conditions.

The last three hypotheses were tested in an attempt to identify characteristics of subjects who changed methods of problem solving. The second null hypothesis was that there was no significant differences in the test anxiety scores of the subjects who did not change their methods. The data supported the null hypothesis.

The third hypothesis that there were no significant differences in the intelligence test scores of subjects who changed problem solving methods and the scores of those subjects who made no change was supported by the data.

The last null hypothesis to be tested was that there was no significant difference in the number of girls who changed problem solving methods and the number of boys who changed methods. This hypothesis was supported by the data, and thus it was concluded that boys and girls are equally likely to change problem solving methods.

An attempt was made to identify any trends of changes that occurred in problem solving methods, but no trends in the changes could be identified. Thus it would appear that changes in problem solving methods are highly individualized or contingent upon variables not investigated in this study.

Wednesday, April 19, 1972

8:45 a.m. - 10:15 a.m. Bel Air Room

Research Reporting Section
NCTM Section No. 153

Subject: MATHEMATICS ACHIEVEMENT AND ITS CORRELATES

Presider: Larry L. Hatfield, University of Georgia

Speakers: Larry E. Wheeler, University of Wisconsin, River Falls
THE RELATIONSHIP OF MULTIPLE EMBODIMENTS OF THE REGROUPING CONCEPT TO CHILDREN'S PERFORMANCE IN SOLVING MULTI-DIGIT ADDITION AND SUBTRACTION EXAMPLES

Ronald R. Edwards, Westfield State College
THE PREDICTION OF SUCCESS IN REMEDIAL MATHEMATICS COURSES IN THE PUBLIC COMMUNITY JUNIOR COLLEGE

Irving M. Brauer, Chicago Board of Education
THE EFFECTS OF VARYING DEGREES OF INQUIRY-CENTERED TEACHING ON COGNITIVE ACHIEVEMENT IN MATHEMATICS UNDER DIFFERENT SCHOOL CONDITIONS

Mildred E. Kersh, University of Washington
A STUDY OF MASTERY LEARNING IN ELEMENTARY MATHEMATICS

Reactor: Jon L. Higgins, The Ohio State University

THE RELATIONSHIP OF MULTIPLE EMBODIMENTS OF THE REGROUPING CONCEPT TO CHILDREN'S PERFORMANCE IN SOLVING MULTI-DIGIT ADDITION AND SUBTRACTION EXAMPLES

Larry E. Wheeler
University of Wisconsin
River Falls, Wisconsin 54022

The multiple-embodiment principle, as stated by Z. P. Dienes, asserts that the abstraction of a concept is facilitated by presentation of the concept using a variety of concrete materials. Furthermore, Dienes conjectures that if abstraction is a result of gathering together the common properties of a large variety of situations, it is more likely that the final abstract concept will be applicable to a large variety of applications and situations. Coupled with this, Bruner maintains that the use of concrete multiple embodiments provides the child with a good stock of visual images which will facilitate the child in solving new related problems. Further, there is evidence indicating that low-I.Q. children do not readily benefit from a single embodiment mode of instruction such as the cuisenaire program. Also, some evidence indicates that single-embodiment treatments may hinder the child's capability to generalize his mathematical knowledge.

Based on the above assertions this study sought answers to the following questions:

1. What is the relationship between a child's performance in representing a concept on several concrete embodiments and his performance in applying the concept to new related problems in the symbolic mode?
2. Is the number of embodiments on which a child is able to represent a concept directly proportional to the child's performance in applying the concept?
3. Are the relationships cited in 1 and 2 above consistent for all levels of intelligence?

As a result of this study, several conclusions were noted for pedagogical consideration and for further consideration by educational researchers. First, there has been little applied research pertaining to Dienes' theory of mathematics learning. This is especially true of the multiple-embodiment hypothesis. This study provided empirical evidence establishing the sufficiency of the multiple-embodiment principle in facilitating concept formation. In particular, this study answered questions on the usefulness of multiple embodiments for facilitating concept formation in children of various intellectual abilities.

There is considerable literature emphazing the need for concrete materials in the elementary classroom, yet little empirical research is available giving direction for implementing these materials. With the current trend towards laboratory learning and the cost factor of supplying classrooms with a large variety of materials, research guidance is needed. This study provided classroom guidance for implementing the various concrete materials recommended in teaching the standard algorithms of addition and subtraction.

In order to approach the problem within the context of the elementary school curriculum the regrouping concept for addition and subtraction was utilized. Relative to this concept the following specific goals were established:

1. To analyze the relationship between a child's performance in solving two-digit addition examples involving regrouping on four isomorphic concrete embodiments and his performance in applying the regrouping concept when solving multi-digit addition examples involving regrouping without concrete aids
2. To analyze the relationship between a child's performance in solving two-digit subtraction examples involving regrouping on four isomorphic concrete embodiments and his performance in applying the regrouping concept when solving multi-digit subtraction examples involving regrouping without concrete aids
3. To analyze goals 1 and 2 across three levels of I.Q.

A random sample of 144 second-grade children was categorized according to three levels of abstraction for regrouping by testing their performances in regrouping two-digit addition and subtraction examples on the abacus, the bundling sticks, the place-value chart, and the multi-base arithmetic blocks. The children were grouped into three levels of I.Q. using the Otis-Lennon group I.Q. test. Written tests involving regrouping two-digit addition, multi-digit addition, two-digit subtraction, and multi-digit subtraction examples served as criteria for the analysis. The children tested in this investigation had studied two-digit addition and subtraction examples involving regrouping. The multi-digit examples had not previously been encountered by the children.

A mixed, 3 × 3 factorial design was used to analyze differences between, or interactions among, the means of the three levels of abstraction and three levels of I.Q. Simple correlation coefficients and partial correlation coefficients were computed to analyze the relationship between the number of embodiments a child was able to regroup for addition and subtraction and his score on the multi-digit tests. The children's ages, I.Q.'s and their competencies in the basic number facts were taken in consideration for the analysis.

The analysis indicated there was no significant difference between the means of children in three levels of abstraction for regrouping relative to their performances in solving two-digit addition and subtraction examples in the symbolic mode. However, it was found that children proficient in regrouping two-digit addition and subtraction examples on three or four embodiments scored significantly higher on the multi-digit written tests than children who were not proficient using the concrete materials. This relationship was consistent across all levels of intelligence. Finally, significant correlations were found between the number of embodiments children were able to regroup for two-digit addition and subtraction and their performances on the multi-digit addition and subtraction tests.

It was concluded that children proficient in regrouping two-digit addition and subtraction examples on three or more concrete embodiments possess a significantly higher level of understanding of the regrouping concept than children without this proficiency with the concrete aids. Secondly, there is a significant relationship between the number of concrete embodiments the children in this study were able to manipulate and their performances on multi-digit written addition and subtraction examples, even when age, I.Q. and competence with the basic number facts were held constant. Relative to the regrouping concept, the evidence supports Dienes' multiple-embodiment hypothesis that concept formation is facilitated through the use of a variety of materials.

THE PREDICTION OF SUCCESS IN REMEDIAL MATHEMATICS COURSES IN THE PUBLIC COMMUNITY JUNIOR COLLEGE

Ronald R. Edwards
Westfield State College
Westfield, Massachusetts 06517

In this study data on 181 remedial mathematics students enrolled in seven public community junior colleges were used to develop a regression equation for predicting success in remedial mathematics courses. Appearing as independent variables in the regression equation were: high school average, standard achievement scores in mathematics and English from the CEEB Comparative Guidance and Placement Program, an attitude toward mathematics score, and a mathematics interest score. The criterion variable was a dichotomous variable based on course grades in remedial mathematics. The resulting regression equation was cross-validated using a second sample of 178 remedial students from the seven colleges. Using this regression equation prediction of success in remedial mathematics courses was made correctly 71 percent of the time.

In keeping with their basic philosophy of "higher education for those who can profit from the instruction" and by the necessity of their "open-door" admissions policy, most community junior colleges offer remedial courses in basic skills, including mathematics. In these remedial courses students who have been found deficient in mathematics are given an opportunity to overcome their deficiencies and enter the "regular" mathematics curriculum of the college. Reports of high attrition rates in these courses (5) bring into question the "salvage function" of remediation in the community college. The salvage function was questioned by Birnbaum who suggested that the community college transfer program acts as a filter through which potentially successful baccalaureate candidates can pass, rather than "a program which strengthens marginal students through counseling and remediation" (1: 247). With little research on the effectiveness of remedial programs, institutions have tended to implement courses and programs by trial and error.

To add to the knowledge about remedial programs in mathematics, this study was concerned with the problem of predicting success in remedial mathematics courses in the public community junior college. In other prediction studies in the community college Worsley (8) and Williams (7) attempted to predict overall grade-point average of junior college freshmen. Worsley was able to predict first-semester grade-point average from high school grade-point average and ACT standard scores. However, Williams using similar independent variables was unable to predict grade-point average for C-average high school students entering the junior college. Morgan (4) was able to predict success in introductory mathematics courses at the junior college from four predictor variables: score on the Cooperative Mathematics Test, years in high school mathematics, high school grade-point average in mathematics, and age. Also, Corotto (2) reported significant correlations between success in a beginning (remedial) college algebra course and ACE (quantitative) and ACE (total) scores.

The purpose of this study was to determine those factors from a selected domain (of 10 possible predictors) which were the best predictors of success in remedial mathematics courses in the public community junior college and to develop a regression equation based upon these predictors.

A sample of 359 remedial students enrolled in seven public community colleges in the fall semester (1970) provided the data for the study--six colleges were in Connecticut, one in Massachusetts. In the regression analysis the following independent variables were identified: scores from the CEEB Comparative Guidance and Placement battery (Reading, Sentences, Mathematics, and Mathematics Interest), high school average, number of class hours for which registered, number of credit hours given in remedial mathematics, and attitude toward mathematics score from the Dutton Test (3:28), work status while attending college, and identification of sex. The criterion variable was a dichotomous variable based on grades in remedial mathematics courses--A,B,C, or Satisfactory for success and D,F,W, or Unsatisfactory for failure. The total sample was divided into two random samples (main and cross-validation) stratified by college.

The stepwise multiple regression program from the Tele-Storage and Retrieval, TSAR, System (6) for the IBM System 360 was used in the analysis (facilities of the University of Connecticut Computer Center were used in the analysis). Included in the statistical analysis were: (1) comparison of means for the independent variables for male and female sample data and for successful and unsuccessful sample data; (2) regression analyses for the main, cross-validation, male and female sample data; (3) determination of a regression equation for predicting success in remedial mathematics courses based on the main sample data; (4) determination of the predictive capabilities of the regression equation based on the cross-validation sample data; and (5) a biserial correlation analysis between the criterion variable and scores on a standardized mathematics test (Mathematics Test from the C.G.P. battery) administered at the end of the fall semester--using a sample of 117 subjects.

The Dutton Attitude Test was administered in thirty-four separate testing sessions which were scheduled within a week of the beginning of the fall semester (1970). The total number of students taking the Dutton Test was 621, which represented 78 percent of the total remedial enrollment at the subject colleges. Students' C.G.P. scores and high school averages were obtained from registrara and admission officers. From the list of 621 subjects, 262 were eliminated because of incomplete records. The remaining 359 subjects represented 45 percent of the total remedial enrollment at the subject colleges. Preadmission testing sessions for the C.G.P. bettery were conducted individually at the subject colleges.

At the end of the fall semester the Mathematics section of the C.G.P. battery was administered to a sample of 117 remedial students. These data were used in a correlation study between performance in remedial mathematics and achievement in mathematics. The final step in collecting data was to obtain final grades in remedial mathematics courses for the 359 subjects in the study--a "1" was assigned as a measure of success and "0" as failure.

When the means and standard deviations of the independent variables for the male and female subjects were examined (see Table 1), it was found that the male subjects averaged significantly higher than the female subjects in Working Status, Mathematics Test score, and in Mathematics Interest. Also, males averaged higher, though not significantly, in Attitude Toward Mathematics and Reading. The female subjects significantly surpassed the males in High School Average and score on the Sentence Test. Females averaged higher, though not significantly, in the Number of Class Hours.

When successful remedial subjects were compared with unsuccessful subjects (see Table 2), it was found that the successful group averaged higher on all variables except Work Status. The successful remedial students scored significantly higher on Reading, Sentences, Class Hours, Mathematics, and on their High School Averages.

For the main and cross-validation sample data the multiple correlation coefficients for the ten independent variables were 0.35 and 0.42. Both were significant at the 0.05 level of significance and there was no significant difference between these two multiple R's at the 0.05 level.

To determine the best combination of independent variables to be used in an equation for predicting success in remedial mathematics courses, different orderings of the independent variables were tried in the analysis (starting with the significant predictors noted in Table 3). This was accomplished using the ORDER option in the TSAR Program--this allows the investigator to specify the order in which the independent variables are entered into the regression analysis. The order for entering variables that was found to be best in terms of its predictive capabilities was : High School Average, Mathematics Test score, Attitude Toward Mathematics score, and Mathematics Interest score. The multiple R for these five variables was 0.33 which was significant at the 0.05 level and was not significantly different from the multiple R determined for all ten variables.

The regression equation based on these five variables was:

$$Y = 0.1809\ X_1 + 0.0132\ X_2 + 0.0129\ X_3$$
$$+ 0.0066\ X_4 - 0.0006\ X_5 - 0.6526,$$

TABLE 1

MEAN DIFFERENCES FOR THE INDEPENDENT VARIABLES FOR MALE AND FEMALE GROUPS

Variable**	Mean-Males N=228	Mean-Females N=131	Diff. Means	Computed z-values
DT	5.09	5.01	0.08	0.50
HSA	1.77	2.15	-0.38	-6.33*
RDG	45.19	44.31	0.88	0.93
SEN	43.71	46.81	-3.10	-3.23
MA	42.37	40.42	1.95	3.48*
WP	1.08	0.88	0.20	3.33*
CH	12.75	13.24	-0.49	-1.26
MI	10.82	8.58	2.24	2.45*

*Significant at the 0.05 level.

**The independent variables were: DT - Attitude Toward Mathematics score [1-11 Thurston scale]; WP - Work Status [2, 1, or 0 for full-time, part-time, or no job]; CH - Number of Class Hours; HSA - High School Average [on four-point scale]; RDG - Reading Test score [C.G.P. Tests were scaled with mean 50 and standard deviation 10]-; SEN - Sentence Test score [C.G.P.]; MA - Mathematics Test score [C.G.P.]; Mathematics Interest score [range 0-32].

TABLE 2

MEAN DIFFERENCES FOR THE INDEPENDENT VARIABLES FOR SUCCESSFUL AND UNSUCCESSFUL GROUPS

Variable	Mean for Successful N=217	Mean for Unsuccessful N=142	Diff. Means	Computed z-values
DT	5.16	4.94	0.22	1.37
HSA	2.04	1.70	0.34	5.66*
RDG	45.74	43.45	2.29	2.31*
SEN	46.09	42.76	3.33	3.50*
MA	42.56	40.27	2.29	3.53*
WP	0.98	1.05	-0.07	-1.15
CH	13.16	12.58	0.58	2.00*
MI	10.41	9.45	0.96	1.03

*Significant at the 0.05 level.

TABLE 3

INTERCORRELATIONS OF THE INDEPENDENT VARIABLES AND CRITERION VARIABLE FOR THE TOTAL SAMPLE

MF	.02	.17	-.08	-.31	.05	-.16	.17	.12	-.02	-.08
DT		.06	-.07	.08	-.26	-.25	.07	.50	-.02	.07
WP			-.38	-.12	-.04	-.07	.06	.00	-.00	-.06
CH				.19	.11	.13	.08	-.05	.03	.10
HSA					.09	.23	.04	-.00	.00	.29*
RDG						.65	.38	-.17	.05	.13*
SEN							.35	-.20	-.03	.18*
MA								.08	-.07	.19*
MI									-.05	.05
CR										-.06

*Significant at the 0.05 level.
**The criterion variable is represented by PF.

where the X_1's represent the following independent variables:

X_1	High School Average
X_2	Mathematics Test score
X_3	Attitude Toward Mathematics score
X_4	Sentence Test score
X_5	Mathematics Interest score

How reliable is this regression equation in predicting success within the subject colleges?

When a cut-off for success and failure in remedial mathematics courses was set at 0.60 (for the computed criterion or Y-value) the regression equation led to correct predictions 71 percent of the time (for the cross-validation sample subjects). Lowering the cut-off score slightly decreases the predictive capabilities of the equation (69 percent at 0.55 and 65 percent at 0.50).

Considering subjects with computed Y-values above 0.70, success in remedial mathematics could be predicted correctly 88 percent of the time.

To determine if there existed a significant relationship between students' performance in remedial mathematics and achievement in mathematics, a correlation (biserial) between the criterion scores and scores on the Mathematics section of the C.G.P. Test (administered at the end of the semester) was computed. The computed correlation coefficient was 0.64, which was found to be significantly different from zero at the 0.05 level of significance. Thus, a positive relationship existed between performance in remedial mathematics courses and achievement in mathematics as measured by a standardized test. However, the level of achievement attained (as measured by the standardized test) at the end of the remedial course was in question, since 57 percent of the students taking the test failed to exceed the score recognized by most colleges as the cut-off for placement in remedial mathematics (this score was 48 on the standard test with mean 50 and standard deviation 10). Also, 42 percent of the successful remedial students failed to exceed this cut-off score.

It was concluded that prediction of success in remedial mathematics courses can be made correctly 71 percent of the time using five select predictors: High School Average, Mathematics Test score (C.G.P.), Attitude Toward Mathematics score (Dutton Test), Sentence Test score (C.G.P.), and Mathematics Interest score (C.G.P.). Since 42 percent of the students passing remedial mathematics courses scored less than 48 on the test which "placed" them in remedial mathematics, an additional conclusion must be that either (a) the cut-off score for placement in remedial mathematics was set too high (i.e., too near the mean of 50 relative to the standard deviation of 10)*, or (b) the criterion for determining that a student needs to take remedial mathematics does not really reflect what the institution wants and therefore should be replaced by an instrument which does, or (c) the students are passed in remedial mathematics courses without achieving the level of mathematical competence (indicated by a score of 48 on the test) for reasons other than their gain in mathematical achievement.

---

*If the cut-off had been set one standard deviation below the mean (i.e., set at 40), then 55 percent of the sample (N=359) would not have been placed in remedial mathematics courses in the fall semester.

REFERENCES

1. Birnbaum, Robert. "Why Community College Transfer Students Succeed in 4-Year Colleges--The Filter Hypothesis." The Journal of Educational Research, 63 (February, 1970), 247-249.

2. Corotto, Loren V. "The Prediction of Success in Initial College Mathematics Courses." The Journal of Educational Research, 56 (January, 1963), 268-271.

3. Dutton, Wilbur H. "Measuring Attitudes Toward Arithmetic." The Elementary School Journal, 55 (September, 1954), 22-31.

4. Morgan, William P. "Prediction of Success in Junior College Mathematics." The Mathematics Teacher, 43 (March, 1970), 260-263.

5. Roueche, John E. and Hurlburt, Allan S. "The Open-Door College: The Problem of the Low Achiever." Journal of Higher Education, 39 (November, 1968), 453-456.

6. Tele-Storage and Retrieval System: User's Manual. Duke University: Duke University Computation Center, 1969.

7. Williams, R. A. "An Assessment of the Success of C Average High School Graduates in Grand Rapids Junior College." Unpublished Ph.D. dissertation, Michigan State University, 1963.

8. Worsley, R.L. "An Analysis of Selected Variables and the Prediction of Educational Achievement of Junior College Freshmen." Unpublished Ed.D. dissertation, Arizona State University, 1969.

THE EFFECTS OF VARYING DEGREES OF INQUIRY-CENTERED TEACHING ON COGNITIVE ACHIEVEMENT IN MATHEMATICS UNDER DIFFERENT SCHOOL CONDITIONS

Irving M. Brauer
Department of Operations Analysis
Chicago Board of Education
Chicago, Illinois

The purpose of the present study was to investigate the effects of three aspects of the school environment on the relationship between varying degrees of emphasis on inquiry-centered teaching in mathematics classes and cognitive learning outcomes in mathematics. Emphasis on inquiry-centered teaching was conceptualized as a continuum which ranged from teaching procedures that stressed drill and rote learning to those that stressed student participation in the discovery of mathematical concepts and solutions to problems. The three aspects of the school environment considered were (1) school learning climates fostered by teaching procedures of the faculty, (2) the teachers' feelings about the school's restrictions on teaching due to set syllabi, textbooks, methods or examinations and (3) the school population's socio-economic climate as reflected by the types of occupations of the students' fathers.

It was hypothesized that the relationship between the degree of emphasis on inquiry-centered teaching and mathematics cognitive outcomes would be non-linear and could be represented meaningfully as a polynomial. Furthermore, this relationship would vary with the three aspects of the school environment listed above.

It is clear that investigations of the simple inquiry versus non-inquiry dichotomy do not suffice; the effects of teaching by a combination of methods is overlooked if the focus is only on this simple contrast. Yet the comparison of many combinations of methods simultaneously is infeasible. In order to avoid the difficulties inherent in the study of the innumerable combinations of teaching procedures, the approach in this study was to assume that a property or trait, namely the degree to which inquiry is emphasized, is present to a lesser or greater degree in all teaching methods and to study the effects of variations in this trait on mathematics cognitive outcomes.

Moreover, these effects cannot be determined without consideration of the interaction between certain contingency factors, such as the three factors under study herein, and teaching methods which emphasize inquiry-centered activities to varying degrees.

The data used in this study were obtained from the data bank created by the International Project on the Evaluation of

Educational Achievement (IEA) study.

In the present study, total scores and part scores on the IEA mathematics test bettery were used as the measures of cognitive outcomes. Scales and indices derived from a student opinionnaire and from student and teacher questionnaires provided measures of the emphasis on inquiry teaching in mathematics classes and of the three aspects of the school environment.

From IEA data on each of the three U.S. target populations, eighth, tenth and twelfth graders, two groups of students were selected according to the level of mathematics instruction they had reached. With each group is associated three population units--schools, teachers, and students. Only teacher and school units were studied by the author; thus a total of twelve populations was investigated.

The following steps were taken in analyzing the data:

(1) The design of the study was a model I (fixed effect) factorial design. An analysis of variance was performed in which the effect of the Inquiry-Emphasis factor (degree of emphasis on inquiry-centered teaching) was divided into orthogonal polynomial contrasts and the effects of the three environmental factors into simple contrasts.

(2) The results of the analysis were used to construct regression models and to compute regression coefficients.

(3) These equations were the basis for comparing the various populations for common trends. In view of the large number of significance tests, replication of similar results was important. As criteria of importance it was assumed that if a non-linear polynomial function is to be a meaningful representation, the coefficient of the second degree term or higher must be different from zero at the .05 level or statistical significance, the trend depicting the relationship must represent a rational and consistent extension of present knowledge, and the trend must be similar to that of at least one other population or be part of a pattern of trends among several populations.

Before proceeding into the specific findings of the study, it should be ntoed that the five levels of the Inquiry-Emphasis factor do not describe absolute amounts of emphasis on inquiry-centered teaching and can only be considered relative to each other. For the purposes of simplicity the extreme levels will be designated as "high" and "low." The three intermediate levels are designated as "moderately high," "middle," and "moderately low."

The same reservation also applies to the School Learning Characteristics factor where the three levels are designated as "authoritarian-based," "neutral," and "inquiry-centered."

The two populations for which the most important findings appeared had identical effects which were statistically significant--a population in which the criterion variable was the mean of the arithmetic part scores of each teacher's participating eighth-grade students and a population in which the criterion variable was the mean of total test scores of each school's participating tenth-grade students. For these populations, the quartic Inquiry-Emphasis main effect component and the linear Inquiry-Emphasis by School Learning Characteristics interaction component were significant at the .05 level.

It remained to show that at each level of the School Learning Characteristics factor the trends were similar for both populations.

The form of the trends was then explored by means of regression equations in which the expected mean values of the mathematics achievement scores were expressed as polynomial functions of the Inquiry-Emphasis scores.

Since no socio-economic interactions were significant, only three equations were needed, one for each School Learning Characteristics level, which were obtained by averaging the estimated values over Socio-economic levels.

From the analysis of these equations and the observed and expected means, the following generalizations are suggested for eighth-grade general mathematics and tenth-grade intermediate algebra or plane geometry courses in schools accepting IEA test items as representative of what they teach. In authoritanian-based school-learning atmospheres, approaches high or moderately high in emphasis on inquiry-centered teaching are not as succesful as those which rely on exposition and set problem-solving procedures. Conversely, in schools tending toward inquiry teaching, and to some degree in "neutral" schools, the best strategy was a moderate emphasis on inquiry-centered teaching. For the most part, any extreme position on the degree of emphasis on the inquiry-centered continuum, i.e., a "high" or "low" score, was not the best strategy as opposed to the adjacent moderate approach.

There is another facet of these findings which does not rely on the validity of the students' description of the environment. Scores on the Degree of Emphasis on Inquiry-centered Teaching and the School Learning Characteristics scales can be considered as measures of students' perceptions of teachers' behavior apart from the correctness of such perceptions. Then the previous conclusions can be interpreted as follows:

In schools where students view school as dominantly rote learning, lecturing and drill, students do better with mathematics teachers whom they perceive as acting consistently with this outlook and not as well with those teachers whom they perceive as expecting them to discover things for themselves. Conversely, in schools where students do not perceive school learning as predominantly rote learning and listening, students of mathematics teachers who are viewed as using a moderate amount of inquiry-centered teaching are more effective than those teachers who are viewed as having a tendency toward lecturing and set problem-solving procedures.

Even though there was some evidence that the socio-economic level of the school affects the relationship between Inquiry-Emphasis and mathematics achievement, it was conflicting in nature and therefore no conclusions could be drawn.

There was no evidence that teachers' feelings about school restrictions placed on their teaching affected the relationship between Inquiry-Emphasis and mathematics achievement.

The results of this study emphasize the importance of the behavioral aspect of the educational experience and the need for joint rather than separate departmental decisions about which student behaviors should be elicited. Furthermore, the results imply that teachers who wish to depart from the normal routines and methods that prevail in the school should be especially attentive to those students who do not know what is expected of them in the new program. Finally, these results, once more, indicate the need to control for the effects of school environmental factors in methods research.

A STUDY OF MASTERY LEARNING IN ELEMENTARY MATHEMATICS

Mildred E. Kersh
University of Washington
Seattle, Washington 98105

Is it possible to devise an instructional strategy which brings most students in a class to a mastery level of achievement? Moreover, would the strategy be effective when used in the classes of two different socio-economic communities? The present study investigates these questions with a sample of fifth grade mathematics students using an instructional strategy for mastery learning proposed by Benjamin S. Bloom.[1]

The study is one of a group of studies examining the effectiveness of Bloom's strategy for mastery learning.[2] The studies range in grade level from elementary school to graduate courses and include the content areas of mathematics, English, testing, and measurement, statistics, and psychology.

Bloom's strategy is one implementation of a Model of Mastery Learning suggested by John S. Carroll.[3] Carroll's model states that for a given individual and a given educational objective or task, the amount of learning is a function of the individual's basic aptitude for learning the task, his willingness to engage actively in learning for the amount of time necessary to achieve the criterion of performance, the quality of the instruction given him, his ability to understand this instruction, and the amount of time he is allowed to have for learning. Time is a common term of the five variables defined in the Carroll model.

Implied in Carroll's model are two hypotheses. (A) The degree of learning of a given individual on a given task is a function of the time he actually spends and the time he needs to spend to learn the task. In terms of a function, the statement becomes:

$$\text{Degree of learning} = f\left(\frac{\text{time spent in learning}}{\text{time needed to learn}}\right)$$

---

[1] Benjamin S. Bloom, "Learning for Mastery," Evaluation Comment. May, 1968, pp. 7-12.

[2] James H. Block (ed.), Mastery Learning Theory and Practice (Holt, Rinehart and Winston, 1971).

[3] John S. Carroll, "A Model of School Learning," Teachers College Record. May, 1963, pp. 723-33.

The second hypothesis implied is (B) Everyone can learn to mastery level if he spends the amount of time he needs to master the task. Evidence supporting these two hypotheses is now available from the Mastery Learning Studies.

In Bloom's strategy an attempt is made to alter the amount of time a student spends in studying a task and thus bring about a level of learning determined to be mastery of the task. The sequential steps in the strategy are specification of objectives, designation of mastery score for summative evaluation, unit teaching, formative tests, immediate feedback, diagnosis of errors, alternative procedures, re-test and summative evaluation.

Design

The collection of data to test the hypothesis of the study extended over a period of sixteen months, beginning with the selection of communities for the study and assignment of experimental and control groups.

Each experimental and control group was composed of three classroom units, the class being the sample unit, and the lower SES experimental cell having only two classes. Teachers were selected from volunteers and were randomly assigned to experimental or control classes. A factorial design was used to analyze the main effects of treatment, the main effects of socio-economic class, and the interaction between these factors.

Statement of Hypotheses

To answer the major questions of the study, several hypotheses were proposed.

Hypothesis 1: For pupils of a given socio-economic status, the mean score of the experimental class will be significantly greater than that of two types of control classes on three achievement measures and two measures of attitude.

A middle class community and a lower class community provided students for this study. Pretest scores on three achievement tests (Stanford Application, Stanford Concepts, and Mastery Test) were the covariates in a multivariate analysis of covariance on mean test scores. The class was the sample unit. Control classes within each community were composed of students not engaged in the study as well as those classes which the experimental teachers had instructed in the previous year. The Dutton Attitude Toward Arithmetic with its accompanying attitude continuum were used as attitude measures.

Hypothesis 2: The strategy for mastery learning will be effective in reducing the deficit in mathematics achievement existing in the lower socio-economic classes in relation to the middle socio-economic classes in fifth grade mathematics.

Hypothesis 3: Students who attain mastery on the first administration of the formative diagnostic test for a unit will spend less time on mathematics than those who do not attain mastery on the first administration.

Hypothesis 4: Students who attain mastery on the second administration of the diagnostic test, but not the first administration (Group III), will have increased the time spent on mathematics and will have used the alternative resources in the suggested manner.

Hypothesis 5: Students who do not attain mastery on either administration of a diagnostic test (Group IV) will not have increased the time spent on mathematics nor will they have used the alternative resources in the suggested manner.

Sub-hypothesis: Students who do not achieve mastery on the first three diagnostic tests will not achieve mastery on the last four diagnostic tests, nor on the final examination.

Results

Hypotheses 1, and 2:

The main effects of treatment and socio-economic class showed no significantly different mean scores in either achievement or attitude variables. However, in the middle class community the experimental classes did attain mastery of the Mastery Criterion Test with mean significantly higher at the .01 level when compared with the previous year controls and retained this position through summer vacation.

The remaining hypotheses are concerned with sub-groups which developed within each class. Four groups were identified on the basis of performance on the formative test and re-test. Group I was composed of students who achieved mastery, 85% or better, on the first administration of a diagnostic test and maintained that level of mastery when the same test was again administered some days later; Group II were mastery then non-mastery students; Group III were non-mastery then mastery students; Group IV students failed to attain mastery on either test.

Hypotheses 3 and 4:

Mastery students report spending more time studying mathematics, contrary to the prediction of Hypothesis 4. The evidence does seem to indicate that a movement to Group III is associated with an increase in time spent on mathematics. The pattern from Group II to III on succeeding tests was significantly associated (.03) with an increase in time spent. Group III students did make use of the alternative materials in the prescribed manner accumulating a mean use ratio (number of prescribed materials used in relation to the number of errors on test) of .80 for the middle class community and .70 for the lower class community.

Hypothesis 5:

A movement from I, II or III to Group IV was always associated with a substantial decrease in time spent on mathematics. The materials use ratio for Group IV never exceeded .50 and was usually well below this for most units. This indicated that Group IV students either did not use the alternative materials, or did not use them in the prescribed manner.

Twenty-nine students were identified as failing the first three formative tests. None of the twenty-nine achieved mastery on any three of the last four tests or on the mastery criterion test.

A further investigation was made into the factors that could identify those who attain mastery on the final mastery criterion test. A class achieving 63% mastery on the mastery criterion test was used as the sample. This class differed significantly (.01 level) from the other experimental classes in the allocation of time spent in grouping for instruction. The teacher of the 63% mastery class divided the instructor role between whole class instruction (59%) and individual tutorials (41%). The buddy system, matching a mastery and a non-mastery student for cooperative mathematics work, was used in the 63% mastery class. There were significant differences in the IQ and attitude of Group III and Group IV students, with Group III scoring higher on both measures. Mastery students used printed material to remediate errors on the diagnostic tests, while non-mastery students showed no preference for a particular kind of alternative material.

Conclusions

Some reasons for the lack of significant main effects may be sought by reviewing the characteristics of the student samples of the study and the implication of these characteristics for changes in the strategy used in the experimental classes.

Non-mastery students randomly sampled visual, printed, and personal interaction modes to remediate errors. Perhaps some means could be devised to ascertain the perceptual mode or combination of modes which provide the best quality of instruction for a particular student, providing a guide for choosing alternative materials for auto-instruction.

In the lower class community 21.8% of the students had repeated at least one grade. That a pattern of failure is established by fifth grade may mean that by grade five it is too late to introduce a mastery learning strategy. Since the failing students do not avail themselves of the materials in the strategy, (Use ration $< .27$) they take part in the study only minimally. The year's instruction for this group, except for taking the formative tests, was no different from that received by all students in the control classes.

Extension of a personal interview with non-mastery students to three or more units could convince these students that the strategy does allow some control over their own learning and thus produce a change in attitude toward the view that man is capable of controlling his environment.

Formative tests should be frequent, at least every two weeks. Two or more mastery criterion periods might be adopted for elementary students.

Wednesday, April 19, 1972

3:45 p.m. - 5:15 p.m. Bel Air Room

Research Reporting Section
NCTM Section No. 217

Subject: TEACHER EDUCATION AND EVALUATION

Presider: Jon L. Higgins, The Ohio State University

Speakers: Nicholas A. Branca, Stanford University
AN EXPLORATORY STUDY OF EXPERIENCED MATHEMATICS TEACHERS' STRATEGIES IN LEARNING A MATHEMATICAL STRUCTURE

Clinton A. Erb, University of Vermont
A FORMATIVE EVALUATION OF AN EXPERIMENTAL TEACHER EDUCATION PROJECT FOR JUNIORS IN MATHEMATICS EDUCATION

Ellen L. Bortz, Montgomery County Public Schools
THE EFFECTIVENESS OF VIDEOTAPE AND AUDIOTAPE IN IMPROVING THE COGNITIVE LEVEL OF QUESTIONS IN THE MATHEMATICS CLASSROOM

W. D. McKillip and M. Mahaffey, The University of Georgia
EVALUATING ELEMENTARY SCHOOL TEACHERS KNOWLEDGE OF THE CONTENT AND METHODS OF TEACHING ELEMENTARY SCHOOL MATHEMATICS

Reactor: Larry Sowder, Northern Illinois University

AN EXPLORATORY STUDY OF EXPERIENCED MATHEMATICS TEACHERS' STRATEGIES IN LEARNING A MATHEMATICAL STRUCTURE

Nicholas A. Branca
Stanford University
Stanford, California 94305

A recent series of experiments by Dienes and Jeeves (1965, 1970) included empirical observations of subjects learning mathematical structures embodied in game playing situations. Strategies based on the moves subjects made were defined, a strategy scoring system was devised, and evidence was presented indicating that a subject's retrospective account of how the game worked (termed an evaluation) reflected the moves he had made. Branca (1971) in replicating and extending the experiments of Dienes and Jeeves did not find any relationship between evaluations given and strategies used. In further analysis of the data, Branca and Kilpatrick (in press) conjectured that the strategy scores as Dienes and Jeeves defined them were insensitive to the strategies subjects may have been using. Subsequent studies were suggested that would either revise the scoring system or change the game playing situation so that subjects would not be constrained at each move by the outcome of the preceding move. The purpose of the present study was to explore both of these possibilities.

Each of thirty-six experienced mathematics teachers was given an experimental task in a game playing situation. The task consisted of a set of four elements on which a closed binary operation having the properties of the mathematical Klein group was defined. During the task each play consisted of combining two of the elements of the set resulting in the outcome of the binary operation in the system. The goal was to learn the rules of the game so as to make correct predictions about the outcome of each combination. The subjects were randomly divided into two equal groups. Group A played according to the method initiated by Dienes and Jeeves where the first element of the combined pair was always the element that was the result of the previous binary operation. Group B played having freedom of choice over the first element as well as the second element to be combined. The experimenter kept track of the subject's moves in learning the game, the predictions he made, and the evaluations he gave of how the game worked. When the game was completed, the subject was asked to explain how he want about playing the game. The experimenter recorded the subject's recollection of the moves he made and the strategies he was using.

In an initial analysis of the data, Dienes and Jeeves' scheme was used to classify subjects' evaluations of the game. Most of the subjects regarded the game as being composed of patterns and many of the mathematical properties of the Klein group were discovered. The original strategy scoring system was used on the sequences of subjects' moves. As in the earlier replication, results failed to indicate a relationship between evaluations and strategy scores. A comparison of the data of both groups indicated that Group B subjects were more successful in learning the rules of the game, took fewer trials in doing so, relied less on pure memory in recalling the rules, saw more of the mathematical properties of the game, and were more aware of regularities in how they were playing. A strategy scoring system is presently being devised to relate Group B subjects' retrospective accounts of how they played and their actual sequence of moves.

Although the nature and size of the sample restricts generalization, the conjecture that the original strategy scoring system is insensitive to subjects' strategies was supported. Also, the free choice method in performing the task to learn the rules of the game proved to be much more efficient for the subjects and lends itself more to use as a research tool in analyzing strategies in learning mathematics structures.

REFERENCES

Branca, N.A. Strategies in learning mathematical structures. (Doctoral dissertation, Teachers College, Columbia University) Ann Arbor, Mich.: University Microfilms, 1971. No. 70-26, 765.

Branca, N.A. & Kilpatrick, J. "The Consistency of Strategies in the Learning of Mathematical Structures." Journal for Research in Mathematics Education, 1972, 3, in press.

Dienes, Z. P., & Jeeves, M.A. Thinking in structures. London: Hutchinson Educational, 1965.

Dienes, Z. P., & Jeeves, M.A. The effects of structural relations on transfer. London: Hutchinson Educational, 1970.

A FORMATIVE EVALUATION OF AN EXPERIMENTAL TEACHER EDUCATION PROJECT FOR JUNIORS IN MATHEMATICS EDUCATION

Clinton A. Erb
University of Vermont
Burlington, Vermont 05401

This study was the initial evaluation of an experimental teacher education project for juniors in mathematics education at The Ohio State University. The project was conducted in five junior high schools during the fall quarter and five elementary schools during the winter quarter of the 1970-71 academic school year. This study had three purposes: 1) the identification and evaluation of significant variables related to the juniors, 2) the investigation of the effect of tutoring upon the achievement of eighth graders, and 3) the determination of relationships between selected teacher variables and success in tutoring.

During the first quarter seventy-two participants spent approximately three class periods, twice a week, in a junior high school. Each tutored an eighth grade student in mathematics for one period and was free to explore the school for the other two periods. The eighth graders who were tutored were randomly chosen from a list of students prepared by the eighth grade mathematics teachers. Those eighth graders on the list who were not tutored formed the control group. Concurrent with the school experience each project member attended a seminar twice a week. Some of the tutoring sessions were videotaped and reviewed during the seminar.

During the second quarter sixty-one juniors spent approximately two and one-half hours, twice a week, in an elementary school. Each was assigned to a particular teacher and served as a teaching assistant. Their exact duties were determined by the co-operating teacher and the junior. All participants attended a seminar twice a week.

The participants were measured at the beginning and end of the first quarter and the end of the second quarter. The instruments administered at these times were the Teaching Situation Reaction Test (Duncan and Hough), Index of Adjustment and Values (Robert Bills), Mathematics Teaching Inventory (experimentor designed), and three questionnaires. Massie's Contemporary Mathematics Test for Teachers was administered at the beginning of the first quarter. The eighth graders were tested at the beginning and end of the first quarter. They were given the Comprehensive Tests of Basic Skills, Arithmetic Section, and an Attitude Inventory constructed using selected items from the NLSMA studies. The elementary teachers completed a questionnaire at the end of the second quarter. During the first quarter each project member kept a log of his activities in the junior high school.

Each hypothesis concerning significant differences in mean scores and variances of the project members was tested using a t-test for correlated means and variances. Each variable was evaluated over the entire project and each quarter individually. The hypotheses concerning differences between the experimental and control groups of eighth graders were tested using an analysis of covariance. Pre-test measures were used as covariates for post-test measures. A correlation matrix of fifty-three identifiable variables was computed.

The results of the analyses for the project members revealed a significant increase in the Teaching Situation Reaction Test for the second quarter and the entire project. The participants also increased significantly on the Composite score, Perception of Student-Teacher Roles, and Strategies of Teaching Mathematics subscales of the Mathematics Teaching Inventory for both quarters and the entire project. The participant's measure of self-concept on the Index of Adjustment and Values increased steadily over the two quarters but did not reach significance. There was no significant difference in variance for any of the measures.

The tutored group of eighth graders scored significantly higher than the non-tutored group on the Math vs. Non-Math and Facilitating Anxiety subscales of the Attitude Inventory. The tutored group had a greater increase in computation scores but the difference between the experimental and control groups was not significant.

There were no significant correlations between the project member's pre-first quarter measures and success in tutoring as defined by gain scores in computation and the subscales of the Attitude Inventory.

The responses to the questionnaires added another source of data. The desire to work in the schools and participate in "practical and relevant" experiences was quoted by two-thirds of the project members as the reason for choosing the project over the "traditional" program. Tutoring and videotaping were the two features of the first quarter cited as contributing the most to the project.

The exposure to teachers and students was listed by half of the juniors as the most beneficial part of the program. The greatest change, according to half of the responses, was in their view of teaching. The participants discovered that there is more to teaching than they previously thought. The change they perceived in themselves as a result of the elementary school experience was taking a personal interest in children and respecting them as individuals.

One activity many felt should be changed was the amount of unstructured time during the first quarter. They felt this time was beneficial at first but became boring and worthless.

In response to a question of what should definitely not be changed, slightly less than one-half said the school experience and, in particular, the work in the elementary school should remain. Tutoring and the use of the videotape recorder were also mentioned by more than twenty-five percent of the participants.

On the questionnaire for the elementary school teachers, slightly over eighty percent said that the junior freed her to do things she might not have been able to do otherwise. All but four teachers said the project members were a help in the classroom. Concerning the strengths of the project, one-third of the teachers not only saw it as aiding the college students but also as benefiting the faculty and students of the participating schools. Another strength listed by almost two-thirds of the teachers was the practical experience gained in handling a classroom.

The results seem to indicate that the experiences provided a realistic framework and point of reference from which the project members could apply the theoretical content of their courses. The results of the tutoring seem to indicate that the eighth grade students acquired a better image of mathematics and greater self-confidence. The responses from the questionnaires seem to relate a positive reaction towards the program from both the participants and cooperating teachers.

THE EFFECTIVENESS OF VIDEOTAPE AND AUDIOTAPE IN IMPROVING THE COGNITIVE LEVEL OF QUESTIONS IN THE MATHEMATICS CLASSROOM

Ellen L. Bortz
Montgomery County Public Schools
Rockville, Maryland

A purposeful communication between teacher and pupil that leads to growth in the pupil's power to think and to learn is a primary goal of good teaching. One of the most common patterns in pupil-teacher classroom communication is that of Teacher Question . . . Pupil Response . . . Teacher Reaction. Research indicates that the type of question asked by the teacher strongly influences the type of response from the pupil. It is reasonable to believe that the pupil's response reflects the level of his inner thinking, and that the thought processes involved in the search for answers are the processes he will acquire.

When the pupil's response is incomplete in some way, the teacher's reaction, in the form of a follow-up question, can further influence the pupil's thinking on the subject. The ability to deviate from a plan to pursue an idea further with follow-up questions is regarded as desirable. It is even more desirable if the questions encourage a higher level of thinking in the pupil.

Thus, questioning behaviors to be encouraged in the student teacher include the ability to formulate questions of high cognitive level, and to ask follow-up questions, of as high a level as possible, as the situation indicates.

Two factors of recent origin--one cognitive, the other electronic--affect the choice of procedures for the development of questioning ability in the pre-service student. The cognitive factor is the development of systems for classifying questions according to cognitive level. The electronic factor is the increasing availability and accessibility of the recording media--videotape and audiotape--for use as recall aids. An evaluation of the usefulness of these media for the improvement of the questioning ability of student teachers, as measured by a question classification system, is timely.

Thirty-six student teachers--eighteen elementary and eighteen at the secondary level--were videotaped individually in microteaching sessions emphasizing questioning. The topic was "Quadrilaterals" at the upper elementary and secondary levels, and "Shapes" at the primary level.

The student teachers were then given a self-instructional unit which included printed material specially written for this project. This was followed by a practice microteaching session which was used for a self-evaluation. One group (Video) was allowed to watch the videotaped practice session for the self-evaluation segment; a second group (Audio) listened to, but did not see, the practice session; and the third group (Self) performed the self-evaluation without recourse to either the video or the audio protion of the taped practice session. A final tape was made of microteaching on the same topic, but with different students, as the first session.

A team of raters classified the questions on the first and third tapes as either Recall or Higher Order, and also as either Probing or Teacher-Initiated. The percentages of Higher Order and of Probing questions rated were then calculated, and the differences between the percentages of Tapes I and III were evaluated statistically.

One Way Analysis of Variance tests revealed no significant difference between groups at the .10 level:

1. In the increase of percentages of Higher Order questions between the three groups
2. In the increase of percentages of questions which were both Higher Order and Probing.

At the .10 level, a One Way Analysis of Variance Test led to a rejection of the null hypothesis of no difference between the groups in the Probing increases. The Tukey test led to a rejection of the null hypothesis of no difference between the Audio and Self groups.

No correlation was found between the gain scores for Higher Order and Probing questions.

The results of this study indicate that the use of videotape for the purposes of self-critique is not of such significant benefit as to recommend its use in the development of the ability to ask Higher Order questions. On the other hand, there is an indication that audiotape self-critique may be of help in the development of the ability to ask Probing questions fewer in number, but of higher cognitive level.

EVALUATING ELEMENTARY SCHOOL TEACHERS KNOWLEDGE OF THE CONTENT AND METHODS OF TEACHING ELEMENTARY SCHOOL MATHEMATICS

W. D. McKillip*, M. Mahaffey*, M. A. Byrne, T. J. Cooney
E. J. Davis, and J. W. Wilson
The University of Georgia
Athens, Georgia 30601

O. R. Brown
Federal City College
Washington, D.C.

During fiscal 1970 the Atlanta, Georgia, Public Schools initiated a program designed to insure that, in the long run, all elementary school teachers meet or exceed "minimum standards of competency" in the teaching of reading and mathematics, and in classroom management. This study is directly concerned only with the teaching of mathematics.

To insure that all elementary school teachers meet or exceed minimum standards of competency in the teaching of mathematics, it is necessary to develop statements of such standards and to devise instructional material through which teachers can attain these competencies. The purposes of this investigation were (1) to locate those content areas and methodological topics in which minimum standards were most needed and to set tentative (and necessarily somewhat vague) standards, (2) to determine the content areas and methodological topics in which teachers fall below the tentative minimum standards and (3) to devise specific competency standards and to create instructional materials to help teachers to meet these standards.

Conceptual Framework: Content.

There have been a number of studies over the last 20 years which, taken all together, indicate that many elementary school teachers are deficient in knowledge of the mathematical content which they are supposed to teach. While it was expected that this general finding would be confirmed, the present study went beyond that general conclusion in three ways. First, the study was focused on a particular well defined group of teachers. Second, the study was designed to determine the areas in which need for knowledge is most seriously different from existing knowledge, thus establishing a priority ordering of topics. Third, the study was designed to contribute to the development of specific statements of minimal competencies needed to teach various topics, going substantially past the general statement that teachers "should be competent in the content of elementary school mathematics."

*To present paper.

Conceptual Framework: Methods.

The long-range goal of this project - the various studies, the development of materials, and the implementation phase - is the improvement of classroom mathematics instruction as evidenced by pupil performance criteria. One component of the comprehensive investigation dealt with methodology and teacher behavior. This study was designed to produce systematic observation and assessment of the knowledge and use of various practices of teaching mathematics within the population of fourth, fifth, and sixth-grade Atlanta teachers.

There is little agreement on techniques for the assessment of mathematics teaching methods or the criteria for acceptable practices. Therefore, several techniques were used: classroom observation, unstructured interviews (teachers, students, principals), indirect observation (class logs), participation in the teaching process, testing sessions, structured interviews, and questionnaires. Flexibility in approach was necessary; the emphasis was on information gathering and organization.

First, a set of desirable classroom practices were selected through literature review, observation, and project staff debate. Second, the study attempted to identify practices from our list that were not at an acceptable level of proficiency by teachers in the population. This in effect aided us in establishing a priority ordering of topics. Third, the study attempted to develop specific statements of minimal competencies needed for teaching mathematics in the elementary school.

Methodology is but one component of the teacher education process and in the development of instructional materials. We are committed to the concurrent development of competencies in content and methods. We recognize competencies in mathematics as necessary attributes of a good teacher and are aware of the literature demonstrating that elementary teachers are lacking in mathematics competencies. Many teachers also lack ability and knowledge of fundamental principles of teaching mathematics, and even where both knowledge of content and knowledge of methodology are present, they may not be evident in teachers' classroom behavior.

Procedures and Analysis

Information was collected in four ways. (1) Four schools, selected so as to reflect the range of ethnic and socio-economic conditions existing in Atlanta, were opened to the investigators. Approximately 25 fourth, fifth and sixth grade teachers in these schools were visited by at least one member of the project staff. The investigator observed the teaching of mathematics and usually the teaching of other subjects also, and on occasion taught mathematics lessons to invidivuals or groups. A staff member generally stayed with a teacher for half a day. The teachers and principals were interviewed informally during these visits. (2) Tests covering methods of teaching elementary mathematics and elementary mathematics content were administered to a sample of 71 teachers. The following general topics were included in the methods test: Diagnosis, explanatory task, selection of appropriate models, diagramatic illustration, instructional aids, and knowledge of basic learning theory. The following general areas were included in the content test: Arithmetical structure, numeration, word problems, geometry, measurement, rational numbers, whole numbers, and number theory.

The test was at the level, approximately, of existing 6th and 7th grade textbooks. Personal data were also collected from these teachers at the time of testing. (3) The teachers kept a log of mathematics teaching activities for one week and returned it to the investigators. (4) Structured interviews concentrating on problems and procedures for teaching mathematics were held with 35 of the 71 teachers participating in the testing sessions.

The information collected in classroom observation and interviews was summarized on an anecdotal basis. The information from the methods test was scored 0, 1, or 2. Answers that were wrong, included no explanation, or included one method when several were asked for, were scored 0. Answers that were correct and adequate were scored 1. Answers that showed diversity and creativity in explanations were scored 2. The teachers were then assigned a total score average for comparison purposes. Responses to the interviews were summarized both anecdotally and descriptively. Personal data (age, length of service, degreed held, teaching assignment, courses on methods of teaching mathematics, etc.) were used to establish subgroups of teachers for informal comparisons.

Information obtained from the mathematics content test was analyzed on an item-by-item basis and also on a topic subscore basis. Total scores were considered in comparing the performance of subgroups within the sample.

The observations and impressions formed by the investigators were collected in an anecdotal sense.

a. Many teachers do not or cannot clarify the ideas they present by means of a picture or diagram where this would be obviously appropriate.
b. Drill activities are too slow and too dull to be effective. Several more specific observations on this point will be included.
c. Teachers make little use of standard audio-visual devices; the chalkboard is, however, used rather effectively by most teachers.
d. The teachers were, in general, highly motivated. They used available resources to the limit of their knowledge and ability. Available resources were, however, in short supply.
e. The teachers were not effective "explainers." The rationale offered for a procedure was virtually always related to "How" rather than "Why." Information tended to be treated as a collection of facts, true but mysterious, which have to be learned but can never be understood.
f. Teachers, not infrequently, communicate false information to their students.

Results of the Content Test.

Interpretation of total scores and subscores indicate the following trends.

a. Teachers performed better on items which are traditional in elementary school mathematics than on items which have recently been introduced.
b. Teachers who report having taken more college mathematics courses, graduate and undergraduate, scored better than those teachers who report fewer courses.

c. Younger teachers scored better than older teachers.
d. Teachers who teach mathematics scored better than those who do not but,
e. Teachers who teach mathematics in a departmentalized organization scored no better than those who teach all subjects including mathematics in a self-contained classroom.
f. The influence of previous mathematical preparation, possibly reaching back to the teachers' own elementary school background, seems highly significant.

Results of Methods Test

The following are some general results obtained from the methods tests. More specific data will be included in a report to be distributed at the time of the presentation.

a. Special attention should be given to teachers who finished a Bachelor's Degree 10 to 20 years ago. (Age 30-40 years)
b. Teachers who finished the Bachelor's Degrees over 20 years ago and who are over 40 scored better on the methods test than teachers in the 30 to 40 age range.
c. In the items requiring explanation many teachers reverted to "old and tried" methods. For example, frequent responses were "invert and multiply" and "line up the decimals when adding."
d. A definite weakness in diagnostic skills was evident.
e. In general, the teachers seemed unable to relate mathematics and the real world by finding concrete examples of mathematical concepts or processes.

Conclusions

Even though the sample was very likely biased in an undetermined direction, the range of scores on the sample show the extreme variation between individuals in the population. When we take into account the results of the methods test and personal data it appears that the "typical," short range, inservice programs are not sufficient to develop and maintain the teachers' methodological skills. Also, the quality of undergraduate programs seems to be highly variable.

Looking at the results of the methods test in conjunction with the observations and interviews, one must conclude that the following areas should receive priority in preparing instruction materials:

a. Materials on a diagnostic approach to teaching.
b. Alternate methods or approaches to the teaching of arithmetic concepts.
c. Techniques of Effective Drill.
d. Examples, lessons, and presentation of arithmetic concepts through the use of concrete materials.
e. Exposure to a multimedia approach to teaching arithmetic.

Instructional materials will ultimately be needed in all content areas of the elementary mathematics curriculum. Even questions as simple as "Which is less, 0.9942 or 1.0013?" were missed, in this case by 5 of the 70 teachers who attempted an answer. Looking only at the results of the content test, one would say that materials were needed in the following priority order.

a. Most urgently needed: Materials on word problems and on geometry.

b. Average need: Arithmetical structure, rational numbers, whole numbers.

c. Less urgent need: Numeration, Measurement.

The information obtained from observation and content testing was used in conjunction with information from other tests and judgments about teaching responsibilities to arrive at decisions regarding subsequent steps for the project. The project staff, in cooperation with the Atlanta Public Schools, is presently engaged in the preparation, trial implementation, and formative evaluation of these instructional materials.